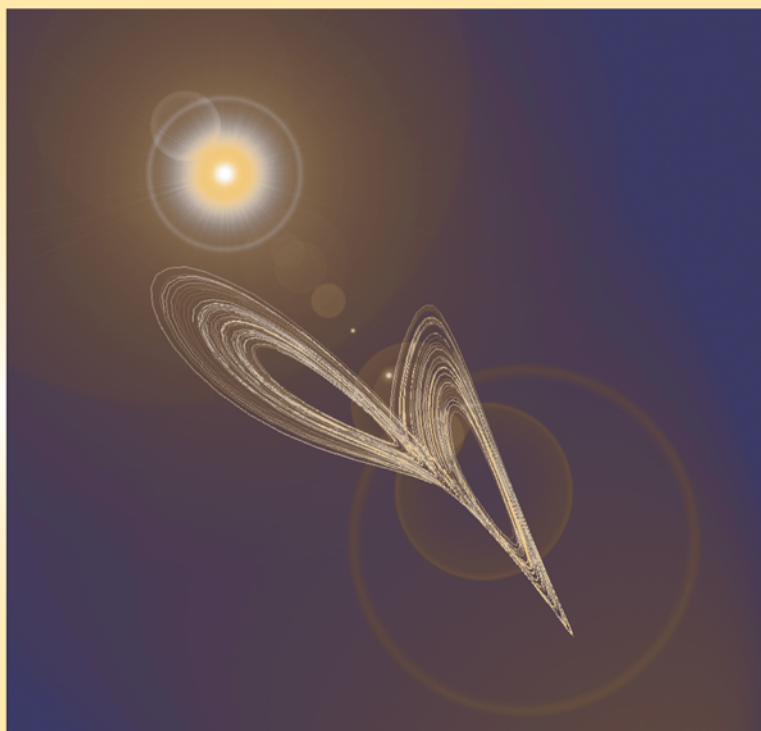


Control Engineering Series

Chaos in Automatic Control



edited by
Wilfrid Perruquetti
Jean-Pierre Barbot

VISIT...

LANZAROTE
Caliente.COM

Chaos in Automatic Control

CONTROL ENGINEERING

A Series of Reference Books and Textbooks

Editor

FRANK L. LEWIS, Ph.D.

Professor

Applied Control Engineering

University of Manchester Institute of Science and Technology

Manchester, United Kingdom

1. Nonlinear Control of Electric Machinery, *Darren M. Dawson, Jun Hu, and Timothy C. Burg*
2. Computational Intelligence in Control Engineering, *Robert E. King*
3. Quantitative Feedback Theory: Fundamentals and Applications, *Constantine H. Houpis and Steven J. Rasmussen*
4. Self-Learning Control of Finite Markov Chains, *A. S. Poznyak, K. Najim, and E. Gómez-Ramírez*
5. Robust Control and Filtering for Time-Delay Systems, *Magdi S. Mahmoud*
6. Classical Feedback Control: With MATLAB, *Boris J. Lurie and Paul J. Enright*
7. Optimal Control of Singularly Perturbed Linear Systems and Applications: High-Accuracy Techniques, *Zoran Gajif and Myo-Taeg Lim*
8. Engineering System Dynamics: A Unified Graph-Centered Approach, *Forbes T. Brown*
9. Advanced Process Identification and Control, *Enso Ikonen and Kaddour Najim*
10. Modern Control Engineering, *P. N. Paraskevopoulos*
11. Sliding Mode Control in Engineering, *edited by Wilfrid Perruquetti and Jean-Pierre Barbot*
12. Actuator Saturation Control, *edited by Vikram Kapila and Karolos M. Grigoriadis*
13. Nonlinear Control Systems, *Zoran Vukić, Ljubomir Kuljača, Dali Donlagić, Sejid Tesnjak*
14. Linear Control System Analysis & Design: Fifth Edition, *John D'Azzo, Constantine H. Houpis and Stuart Sheldon*
15. Robot Manipulator Control: Theory & Practice, Second Edition, *Frank L. Lewis, Darren M. Dawson, and Chaouki Abdallah*
16. Robust Control System Design: Advanced State Space Techniques, Second Edition, *Chia-Chi Tsui*
17. Differentially Flat Systems, *Hebertt Sira-Ramirez and Sunil Kumar Agrawal*
18. Chaos in Automatic Control, *edited by Wilfrid Perruquetti and Jean-Pierre Barbot*

Chaos in Automatic Control

edited by

Wilfrid Perruquetti

*Ecole Centrale de Lille
Villeneuve-d'Ascq Cedex, France*

Jean-Pierre Barbot

*Equipe Commande des Systèmes
Cergy-Pontoise Cedex, France*



Taylor & Francis

Taylor & Francis Group
Boca Raton London New York

A CRC title, part of the Taylor & Francis imprint, a member of the
Taylor & Francis Group, the academic division of T&F Informa plc.

Published in 2006 by
CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2006 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-8247-2653-7 (Hardcover)
International Standard Book Number-13: 978-0-8247-2653-9 (Hardcover)
Library of Congress Card Number 2005050539

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Chaos in automatic control / edited by Wilfrid Perruquetti, Jean-Pierre Barbot.

p. cm. -- (Control engineering (Taylor & Francis))

Includes bibliographical references and index.

ISBN 0-8247-2653-7 (alk. paper)

1. Automatic control. 2. Chaotic behavior in system. I. Perruquetti, Wilfrid. II. Barbot, Jean-Pierre, 1958- III. Series.

TJ213.C468 2005

629.8--dc22

2005050539

informa

Taylor & Francis Group
is the Academic Division of Informa plc.

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Dedication

*This book is dedicated to Valérie, Isabelle, Marius, Rosalie,
Thomas, Tristan, and Baptiste*

Preface

Motivations

Chaotic dynamics, first espoused by the French mathematician Henri Poincaré (1854–1912), has received considerable attention over the years since. In 1963, when simulating a simplified model of convection, Edward Lorenz highlighted its unpredictable nature, great sensitivity to initial conditions, and strange attractors. Well-known qualitative methods for studying nonlinear system models and the notion of bifurcation in the phase plane have been largely inspired from the works of Andronov (first published in 1937). Over the years, chaotic phenomena have been mainly investigated from an analysis point of view. Since 1990, considerable developments have occurred in the control and observation of chaotic systems. A huge number of applications have been proposed in the fields of circuit systems, mechanics, physics, avionics, weather forecasting, and more recently, secure communications and cryptography. Around this time, people started considering these problems and several active researchers in this field combined their efforts, thanks to the support of many French institutions.¹ Several tools on normal forms, bifurcations, and chaos have been presented. An international workshop was organized in Lille (September 2003) with the intention of:

- bringing together researchers from different areas of engineering who were interested in chaotic systems;
- promoting some new concepts of modern control theory dedicated to chaotic systems;
- overviewing some recent developments on chaos control for physical and industrial applications.

After this meeting, it was decided by the contributors to collate and present all the theoretical and pedagogical material in a book. The major goal being to cover advanced topics adopted from the field of automatic control in the specific context of control and observation for chaotic systems. In addition, the aim of the book is to familiarize the control systems community with chaos theory and to equip the specialists of chaotic

¹CNRS, GdR Macs, GRAISyHM, LAGIS, ECE-ENSEA, Ecole Centrale de Lille, and so on.

dynamics with many of the most recent advances in modern control theory. A deterministic point of view (using ordinary differential equations and difference equations) is chosen to tackle chaotic systems: this choice is justified by the extended results available on this subject and by the huge number of application domains. The book is organized as follows:

1. The models addressed are mainly ordinary differential (or difference) equations, but these concepts may be analyzed within the framework of other models such as partial differential equations, delay differential equations, algebra differential equations, and so on.
2. Chaos theory is gradually introduced starting from bifurcation theory. More precisely, our focus is on the stability analysis of control schemes dedicated to chaotic systems; however, some ergodic arguments are involved when dealing with the well-known and efficient Ott–Grebog–York (OGY) method. Nevertheless, in this case we highlight the interest of such an approach and refer to the literature, for more detail, such as the well-known Taken’s theorem.
3. As a consequence of the preceding point, the usefulness of tools from control problems such as analysis (stability, observability, controllability, etc.), and design of controller or observer, are particularly highlighted for chaotic systems.
4. The Poincaré normal form was a starting point for some analytical purposes, but recent and important contributions have been made by Wei Kang and Art Krener for the normal form to some analysis and control design.
5. A possible interpretation of the synchronization problem as an observer design problem comes from a well-known paper by Henk Nijmeijer and Ian Marels. However, there also exist some other historical points of view such as those of the Pecorra and Carol. Some of them are discussed here.
6. As the application domains are varied, for example, aeronautics, biology, chemistry, economy, and so on, we only tackle telecommunication and electrical drive problems.

Obviously, all these choices lead to a non-exhaustive presentation of chaos and some important theoretical parts are only very briefly tackled or mentioned. The ergodic approach, presented very briefly in Chapter 6, uses ergodic arguments to explain not only the well-known control schemes such as the OGY and Pyragas methods but also a new one based on H_∞ concepts. The usefulness and potential of the proposed deterministic methods are undeniable. It is also obvious that future developments of the theories proposed here may be carried out using some ergodic arguments.

Book Overview

To achieve the aforementioned goals, and more specifically to deal with fundamental theoretical backgrounds and an interdisciplinary presentation of emergent methods and applications, we organize the book in three parts.

Part I: Open-Loop Analysis

As the meeting was mainly attended by physicists and control people, we decided, for pedagogical reasons, to present some mathematical background on ordinary differential equations and difference equations to understand the concepts involved in the following chapters. This constitutes the core of open-loop analysis:

- *Chapter 1*: an historical, theoretical overview of the discrete time system (difference equations) is given by C. Mira (one of the pioneers in chaos theory). His background information allows the reader to easily understand analysis and design tools using mathematical descriptions.
- *Chapter 2*: background on ordinary differential equations is presented in this chapter and concerns the notion of solutions, and their qualitative properties (equilibrium points, limit cycle and strange attractor, asymptotic behavior, etc.).
- *Chapter 3*: background on Poincaré normal form is recalled to investigate specific bifurcation phenomenon such as Hopf bifurcation.
- *Chapter 4*: this chapter describes in details the approach of Poincaré using homogeneous and found transformations, generalize, to nonlinear control systems. A variety of usual and canonical forms under the action of nonlinear feedback is presented. Applications to feedforward systems and to symmetries are Geometric aspects of the presented results are discussed.
- *Chapter 5*: some systems inherently have a two time-scale behavior captured by a "singular perturbation" approach. This chapter focuses on the interconnection between chaos and singular perturbation phenomenon.

Part II: Closed-Loop Design

Some problems arising in observation (synchronization and observability bifurcation) and control (chaotic and hyperchaotic control and control

bifurcation) are typically closed-loop-oriented. This is the area covered in this part.

- *Chapter 6:* chaotic systems feature unstable periodic orbits and sensitivity to parameters, initial conditions, and external disturbances. To cope with such behavior, some classical model-independent control methods are recalled (OGY and Pyragas). Then, other strategies such as H_∞ , sliding modes, adaptive control, and energy-based control are presented. Finally, recent methods are reported to deal with the control of hyperchaotic systems.
- *Chapter 7:* to deal with chaos synchronization, polytopic observers design is developed for a special class of chaotic systems through the notion of polyquadratic stability.
- *Chapter 8:* as the Poincaré's theory of normal forms for uncontrolled dynamical systems uses homogeneous transformations, a variety of control system normal forms were derived using extended homogeneous transformations. This chapter reports on a unified framework of these normal forms.
- *Chapter 9:* similar to the previous chapter, observability normal forms are introduced and applied to some synchronization problems.
- *Chapter 10:* for synchronization, observer design in the case of nonlinear system with a linear detectable part is important. The Kazantzis–Kravaris and the Kreisselmeier–Engel methods are compared in this chapter.

Part III: Some Applications

This part covers applications and also presents illustrative examples of chaos-based engineering. They are related to wireless transmissions, optics, power electronics, and cryptography using chaos.

- *Chapter 11:* different modulation schemes that allow the transmission of some information with chaotic carriers are described in the context of microwaves.
- *Chapter 12:* code division multiple access (CDMA) is shown to be closely related to chaos-based encryption. Moreover, a nonlinear delay differential equation related to optics and optoelectronics is shown, which act as chaos generators.
- *Chapter 13:* the appearance of self-sustained oscillations in high-performance AC drives, and in particular in field-oriented control of induction motors, due to the existence of Hopf bifurcations, is discussed.

- *Chapter 14*: Chua's circuit combined with an analog circuit observer designed step by step to obtain a chaos-based secure communications channel.
- *Chapter 15*: this chapter is the discrete counterpart of the previous chapter because it addresses a discrete time cryptography point of view based on the inclusion method (DCCIM). This chapter provides a practical implementation of techniques developed in Chapter 9 based on the use of observability normal forms and observability bifurcation analysis.

Acknowledgments

The authors are indebted to their institutions (LAGIS UMR CNRS 8146, Ecole Centrale de Lille, ECS and ENSEA), and to CNRS, GdR Macs, and GRAISyHM. These institutions provided us with the facilities for organizing an International Workshop at Lille (LISAC 03, September 2003) and, in addition, some of these institutions also provided us with a good environment for the editing and completion of this book. Thanks to Zheng Gang, a Ph.D. student of Professor J.-P. Barbot for his help in the conception of the cover figure.

About the Editors



Wilfrid Perruquetti was born in 1968 at Saint Gilles, France. In 1991, he received his M.Sc. in automatic control from the Institut Industriel du Nord. In 1994, he obtained his Ph.D. in automatic control and then joined the Ecole Centrale de Lille as an Assistant Professor in 1995. Since 2002, after receiving the "Habilitation à Diriger les Recherches" in 2001, he has held a full Professor's position at the same institute. He has published over sixty books, journal articles, and conference papers and is the co-editor with Jean-Pierre Barbot of the book *Sliding Mode Control in Engineering* (Marcel Dekker). He is currently working on stability analysis (including various kinds of stability concepts), stabilization (in particular, finite stabilization), and sliding mode control of nonlinear and delay systems.



Jean-Pierre Barbot was born in 1958 in Paris, France. He is the director of the ECS Laboratory and full professor of control systems at Ecole Nationale Supérieure d'Electronique et de ses Applications (ENSEA), Cergy, France. He received the Agrégation (1985) in electrical engineering from the Ecole Normale Supérieure (ENS) de Cachan, France, and Ph.D. (1989) and "Habilitation à diriger des recherches" (1997) from the University of Paris XI, Orsay, France. He has published hundreds of patents, book chapters, journal and conference papers and is the co-editor with Wilfrid Perruquetti of the book *Sliding Mode Control in Engineering* (Marcel Dekker). He has been a visiting professor at several international universities. He is currently working on chaos synchronization (more particularly, observability normal form), hybrid systems, and sliding mode observer, and his application domains are cryptography and electrical drive.

Contributors

J. Aracil

Escuela Superior de Ingenieros
Universidad de Sevilla
Sevilla, Spain

J.-P. Barbot

Equipe Commande des
Systèmes (ECS)
ENSEA, Cergy Pontoise
France

I. Belmouhoub

Equipe Commande des
Systèmes (ECS)
ENSEA, Cergy Pontoise
France

L. Boutat-Baddas

CRAN-CNRS UMR
UHP, Nancy
France

J. Daafouz

CRAN-CNRS UMR
INPL, Nancy
France

M. Djemai

Equipe Commande des
Systèmes (ECS)
ENSEA, Cergy Pontoise
France

F. Gordillo

Escuela Superior de Ingenieros
Universidad de Sevilla
Sevilla, Spain

J. Guittart

IRCOM CNRS Université
de Limoges IUT
Jules Valles
France

W. Kang

Department of Mathematics
Naval Postgraduate School
Monterey, California, USA

A.J. Krener

Department of Mathematics
University of California
Davis, California, USA

L. Larger

LOPMD-CNRS
Université de
Franche-Comté

L. Laval

Equipe Commande des
Systèmes (ECS)
ENSEA, Cergy Pontoise
France

G. Millerioux

CRAN-CNRS UMR
UHP, Nancy
France

C. Mira

Cellule d'Etude des Systèmes
Non Linéaires et Applications
(CESNLA)

Quint Fonsegrives
France

Istituto di Scienze Economiche
University of Urbino, Italie

J.C. Nallatamby

IRCOM CNRS Université
de Limoges IUT
Jules Valles, France

R. Oritega

Laboratoire des Signaux
et Systèmes
LSS/CNRS/Supélec
Gif Sur Yvette, France

W. Perruquetti

LAGIS-CNRS
Ecole Centrale de Lille
Villeneuve-d'Ascq, France

R. Quere

IRCOM CNRS Université
de Limoges IUT
Jules Valles, France

S. Ramdani

Laboratoire EDM
Université de Montpellier I
Montpellier, France

W. Respondek

Laboratoire de Mathématiques
INSA de Rouen
Mont Saint Aignan, France

F. Sales

Escuela Superior de Ingenieros
Universidad de Sevilla
Sevilla, Spain

I.A. Tall

Department of Mathematics
Natural Science Division
Tougaloo College, Mississippi
USA

R. Tauleigne

Equipe Commande des
Systèmes (ECS)
ENSEA Cergy Pontoise
France

C. Dang Vu-Delcarte

LiMSi-CNRS Université de
Paris-sud VI
Orsay, France

M. Xiao

Department of Mathematics
Southern Illinois University
Carbondale, Illinois, USA

Series Introduction

Many textbooks have been written on control engineering, describing new techniques for controlling systems, or new and better ways of mathematically formulating existing methods to solve the ever-increasing complex problems faced by practicing engineers. However, only a few of these books fully address the application aspects of control engineering. It is the intention of this series to redress this situation.

This series will stress on application issues, and not just the mathematics of control engineering. It will provide text that presents not only new and well-established techniques, but also detailed examples of the application of these methods to the solution of real-world problems. The authors will be chosen from both the academic and the relevant application sectors.

There are already many exciting examples of the application of control techniques in the established fields of electrical, mechanical (including aerospace), and chemical engineering. We only have to look around in today's highly automated society to see the use of advanced robotic techniques in the manufacturing industries, the use of automated control and navigation systems in air and surface transport systems, the increasing use of intelligent control systems in the many artifacts available to the domestic consumer market, and the reliable supply of water, gas, and electrical power to the domestic consumer and to industry. However, there are many challenging problems that could benefit from wider exposure to the applicability of control methodologies, and the systematic system-oriented basis inherent in the application of control techniques.

This series presents books that draw on expertise from both the academic world and the applications domains, and will be useful not only as academically recommended course texts but also as handbooks for practitioners in many application domains. *Chaos in Automatic Control* is another outstanding entry in CRC's Control Engineering Series.

Contents

Part I	Open-Loop Analysis	1
1.	Bifurcation and Chaos in Discrete Models: An Introductory Survey	3
	<i>C. Mira</i>	
2.	Tools for Ordinary Differential Equations Analysis	45
	<i>W. Perruquetti</i>	
3.	Normal Forms and Bifurcations of Vector Fields	101
	<i>C. Dang Vu-Delcarte</i>	
4.	Feedback Equivalence of Nonlinear Control Systems: A Survey on Formal Approach	137
	<i>W. Respondek and I. A. Tall</i>	
5.	Singular Perturbation and Chaos	263
	<i>M. Djemai and S. Ramdani</i>	
Part II	Closed-Loop Design	289
6.	Control of Chaotic and Hyperchaotic Systems	291
	<i>L. Laval</i>	
7.	Polytopic Observers for Synchronization of Chaotic Maps	323
	<i>G. Millérioux and J. Daafouz</i>	
8.	Normal Forms of Nonlinear Control Systems	345
	<i>W. Kang and A. J. Krener</i>	
9.	Observability Bifurcations: Application to Cryptography	377
	<i>J.-P. Barbot, I. Belmouhoub, and L. Boutat-Baddas</i>	
10.	Nonlinear Observer Design for Smooth Systems	411
	<i>A.J. Krener and M. Xiao</i>	

Part III	Some Applications	423
11.	Chaos and Communications	425
	<i>R. Quéré, J. Guittard, and J.C. Nallatamby</i>	
12.	Chaos, Optical Systems, and Application to Cryptography	453
	<i>L. Larger</i>	
13.	Indirect Field-Oriented Control of Induction Motors: A Hopf Bifurcation Analysis	481
	<i>Francisco Gordillo, Francisco Salas, Romeo Ortega, and Javier Aracil</i>	
14.	Implementation of the Chua's Circuit and its Application in the Data Transmission	503
	<i>L. Boutat-Baddas, J.-P. Barbot, and R. Tauleigne</i>	
15.	Synchronization of Discrete-Time Chaotic Systems for Secured Data Transmission	527
	<i>I. Belmouhoub and M. Djemai</i>	
	Appendix A. On Ergodic Theory of Chaos	553
	Index	561

List of Figures

2.1	Unit circle: simulation of (2.13)	54
2.2	Infinite number of solutions to the CP of (2.14)	55
2.3	Euler approximates	56
2.4	Limit cycle of (2.24)	64
2.5	Periodic closed orbit of the Van der Pol oscillator (2.25)	66
2.6	Homoclinic orbit	67
2.7	Heteroclinic orbit	67
2.8	Invariance of $\mathcal{A}$	68
2.9	Stability of the $\mathcal{A}$ and its stability domain $\mathcal{D}_s(\mathcal{A})$	69
2.10	Attractivity of the set $\mathcal{A}$ and its attractivity domain $\mathcal{D}_a(\mathcal{A})$	71
2.11	Equilibrium $(1, 0)$ is attractive and unstable	73
2.12	Poincaré section	87
2.13	Branch of equilibrium	90
2.14	Hopf bifurcation of (2.3)	91
2.15	Saddle-node bifurcation of (2.43)	93
2.16	Saddle-node bifurcation with $\dot{y} = -y$	93
2.17	Transcritical bifurcation of (2.44)	94
2.18	Fork bifurcation of (2.45)	95
2.19	Rössler attractor	96
3.1	Hopf bifurcation diagrams	122
3.2	The Poincaré-Bendixson annulus	122
3.3	$\alpha'(\nu_C) < 0$ and $a_1 < 0$	123
3.4	Bifurcation diagram for the Rössler model	133
3.5	Correspondence between (r, z) and (r, ϕ, z)	134
5.1	The Chua's cubic electronic oscillator	276
5.2	Chua's cubic attractor for $\mu = 2$ (5.30) with initial conditions: $x_0 = 0.5, y_0 = -0.5, z_0 = 1$	277
5.3	Slow manifold M_0 associated to system (5.30)	278
5.4	A representation of the slow manifold M_0 of Chua's cubic system	279
5.5	The global motion of the Chua's cubic system	279
5.6	The chaotic attractor of the HR model obtained by numerical integration for $K = 3.18$ and $\varepsilon = 0.004$	281
5.7	The chaotic temporal evolutions $x(t)$	282
5.8	The chaotic temporal evolutions $y(t)$	283
5.9	The chaotic temporal evolutions $z(t)$	284
6.1	Chaotic trajectory	295
6.2	The Poincaré section and an UPO	296

6.3	Schematic explanation of OGY method	297
6.4	Schematic representation of the Pyragas control scheme	300
6.5	The (linear) H_∞ control scheme	304
6.6	Schematic representation of the YLM method	314
7.1	Message-embedded scheme	334
7.2	(a) Error $\ x_k - \hat{x}_k\ $, (b) error $m_k - \hat{m}_k$, (c) plaintext m_k , and (d) recovered plaintext $\hat{m}_k$	339
7.3	Decoder capture screens: matched and mismatched keys	340
8.1	The configuration of ball and beam system	366
9.1	Bidirectional synchronization	379
9.2	Unidirectional synchronization	380
9.3	Inclusion method	394
9.4	Addition method	394
9.5	The Chua circuit	395
9.6	Observation errors	397
9.7	Rosler map phase portrait	399
9.8	Three steps convergence of signal observation error	402
11.1	Architecture of the master-slave synchronization of two chaotic systems	427
11.2	Phase portrait of the Lorenz system in the chaotic regime for $\sigma = 16$, $b = 4$, and $r = 45.92$	429
11.3	x_1 waveforms corresponding to: (a) synchronization and (b) lack of synchronization	430
11.4	Plots of the error signal versus time for a perfect match of the emitter and receiver parameters	430
11.5	Plots of the magnitude of the error signal in the case of a mismatch of master and slave parameters	431
11.6	Architecture of a feedback-type chaotic synchronization	432
11.7	Example of a non-autonomous synchronization system based on an inverse system	432
11.8	Typical coder-decoder based on an inverse chaotic system	434
11.9	Example of a coherent CSK system	435
11.10	Architecture of a COOK system	435
11.11	Architecture of a DCSK system	436
11.12	General architecture of the chaotic oscillator	437
11.13	Schematic of the VCO used for the chaotic oscillator	438
11.14	Frequency of the oscillations of the VCO versus the control voltage	441
11.15	Amplitude of the oscillations of the VCO versus the control voltage	442
11.16	Chaotic oscillator $T/\tau = 64$ $V_{c0} = -4$ V	442
11.17	Bifurcation diagram of the chaotic oscillator for $\alpha = 314rd$	444
11.18	Bifurcation diagram of the chaotic oscillator for $\alpha = 628rd$	444
11.19	Transient set-up of a two-frequencies steady state regime	445

11.20	Spectrum of the output signal in the chaotic regime	447
11.21	Architecture of the chaotic modulator	447
11.22	Structure of the chaotic receiver (demodulator)	448
11.23	Transmitted and received signals in the case of a perfect match between parameters	448
11.24	Transmitted and received signals in the case of a 10% mismatch between parameters	449
11.25	The BER of the chaotic modulator–demodulator	450
12.1	Typical transmission system using chaos encryption	456
12.2	Bloc diagram of the scalar nonlinear delayed dynamic	459
12.3	Bifurcation diagram calculated from Equation (12.2)	460
12.4	Bloc diagram in the adiabatic approximation situation	461
12.5	Bifurcation diagram calculated from a mapping using $\beta f[\cdot]$	461
12.6	The Ikeda ring cavity	462
12.7	The wavelength chaos generator	464
12.8	Experimental trajectories in time and frequency domain	465
12.9	Experimental bifurcation diagrams	466
12.10	Bloc diagram for chaos replication and decoding	468
12.11	Replication error against parameter mismatch	469
12.12	Set-up of the wavelength chaos receiver–decoder	471
12.13	Experimental traces while encoding and decoding a sine waveform	471
12.14	Electrooptic intensity chaos emitter	474
12.15	Emitter–receiver set-up using chaos in coherence modulation	475
12.16	Basic ECLD set-up	476
12.17	Direct optoelectronic feedback in an SC laser	477
13.1	Transversality condition for a Hopf bifurcation	484
13.2	Supercritical Hopf bifurcation	484
13.3	Representation of Equation (13.14) for $c_1 = 4$, $c_2 = 4$, $c_4 = 1$, $c_5 = 1$, and $u_2^0 = 1$	488
13.4	Root locus for $c_1 = 4$, $c_2 = 4$, $c_4 = 1$, $c_5 = 1$, $u_2^0 = 1$, $k_p = 0.1$, $k_i = 1$ and $\tau_L = 0.3$	489
13.5	Values of κ corresponding to a Hopf bifurcation vs. τ_L	490
13.6	Evolution of x_3 in the four simulations	492
14.1	Additive chaos masking	504
14.2	The chaotic parameter modulation	504
14.3	Chua’s circuit	505
14.4	Current–voltage characteristic of the negative resistance	506
14.5	Elaboration of a negative resistance	506
14.6	The negative resistance with double slope	507
14.7	Real current–voltage characteristic	507
14.8	Complete implementation of the Chua’s circuit	508
14.9	Route to chaos	509

14.10	Parlitz's experience	510
14.11	Double scroll attractor for system (14.2) and system (14.3)	513
14.12	Double scroll attractor for system (14.2) and system (14.4)	513
14.13	Observation error for system (14.2) and system (14.3)	514
14.14	Observation error for system (14.2) and system (14.4)	514
14.15	Double scroll attractor for system (14.2) and system (14.5)	516
14.16	Observation error for system (14.2) and system (14.5)	516
14.17	Double scroll attractor for systems (14.6) and (14.8), when we set $E_s = 0$ on a big neighborhood of the singularity manifold $(x_2 + R_0x_3)$	521
14.18	$x_4, \hat{x}_4, E_s$, and the singularity $(x_2 + R_0x_3)$, when we set $E_s = 0$ on a big neighborhood of the singularity manifold $(x_2 + R_0x_3)$	522
14.19	Double scroll attractor for systems (14.6) and (14.8), when we set $E_s = 0$ on a very small neighborhood of the singularity manifold $(x_2 + R_0x_3)$	522
14.20	$x_4, \hat{x}_4, E_s$, and the singularity $(x_2 + R_0x_3)$, when we set $E_s = 0$ on a very small neighborhood of the singularity manifold $(x_2 + R_0x_3)$	523
14.21	Transmission examples	523
15.1	Phase portrait of Burgers map	529
15.2	Phase portrait of the observer	530
15.3	Observation error dynamics on x_1 and x_2 . A zoom on the first 10 iterations (10 stars)	533
15.4	Phase portrait of the Mandelbrot map for $a = 0.2, b = -0.7,$ $c = 0.8$, and $d = 0.291$	542
15.5	Three-dimensional bifurcations diagram	543
15.6	Bifurcations diagram	544
15.7	Mandelbrot map's phase portrait for $c = 0.95$	544
15.8	Arnold's tongue for the Mandelbrot map	545
15.9	The original and recovered picture	548
15.10	The ciphered picture	548
15.11	An example of ciphering/deciphering a text file by the CCMID	549
15.12	Original text	549
15.13	Ciphered text (Figure 15.12) by the Mandelbrot map (in simple precision)	549

Part I

Open-Loop Analysis

1

Bifurcation and Chaos in Discrete Models: An Introductory Survey

C. Mira

CONTENTS

1.1	Introduction	4
1.2	Chaos and Unpredictability	9
1.3	Generalities on Discrete Models	10
1.3.1	Different Forms of Models	10
1.3.2	Maps Obtained from an ODE by a Poincaré Section	13
1.4	Singularities and Bifurcations Common to Invertible and Noninvertible Maps	15
1.4.1	Singularities and Bifurcations	15
1.4.2	Bifurcation Sets: Normal Forms of Exceptional Critical Cases	17
1.4.3	Singularities Sense when the Map is Obtained from a Poincaré Section	20
1.5	Map Singularities and Bifurcations Specific to Noninvertible Maps	22
1.5.1	Singularities and Bifurcations Induced by Noninvertible Maps	22
1.5.2	Fractal Bifurcations Structure of “Embedded Boxes” Type and Chaotic Behaviors	23
1.5.3	Homoclinic and Heteroclinic Situations: Their Bifurcations	25
1.6	Absorbing Areas, Chaotic Areas, Bifurcations	27
1.6.1	Definitions and Properties	27
1.6.2	Chaotic Areas: Microscopic and Macroscopic Points of View	28
1.7	Results on Basins and their Bifurcations	30
1.8	Map Models with a Vanishing Denominator	31
1.9	Noise and Chaos: Characterization of Chaotic Behaviors	33

1.10 Conclusion	34
References	35

1.1 Introduction

Dynamics is a concise term referring to the study of time-evolving processes. The corresponding system of equations describing this evolution is called a *dynamic system*. *Nonlinear dynamics* is the scientific field concerning the behavior of real systems, linearity being always an approximation. This field, which embraces ordinary differential equations (continuous dynamics) and maps, also called recurrences (discrete dynamics), is too wide to be completely presented in the limited framework of this book. This chapter is limited to an introductory and preparatory knowledge to tackle more complete readings.

Two different approaches have been developed for studying nonlinear dynamics. The first corresponds to *qualitative methods* [9–11]. The “strategy” of these methods can be defined noting that the solutions of equations of nonlinear dynamic systems are in general nonclassical, nontabulated, transcendental functions of mathematical analysis, which are very complex. This strategy is of the same type as the one used for the characterization of a complex variable function by its singularities: zeros, poles, essential singularities. Here, the complex transcendental functions are defined by the singularities of continuous (resp. discrete) dynamical systems such as:

Stationary states which are equilibrium points (resp. fixed points), or periodic solutions, that is, limit cycles in the continuous case (resp. cycles in discrete case); which can be stable or unstable

Trajectories (resp. invariant curves), passing through saddle singularities of two dimensional systems

Stable and unstable manifold for a dimension greater than two

Boundary, or separatrix, of the domain of attraction (or *basin*) of a stable (attractive) stationary state

Homoclinic or heteroclinic singularities (defined subsequently)

More complex singularities of fractal, or nonfractal type

The qualitative methods consider the nature of these singularities in the *phase space* (state space) and their evolutions in the presence of varying system parameters or in the presence of a continuous structure modification of this system (study of the bifurcation sets in a parameter space, or in a function space). Roughly speaking, a *bifurcation* corresponds to a

qualitative change of a system behavior from a very small modification of its parameters or of its structure.

Within the framework of the bifurcation theory for continuous time systems [described by ordinary differential equations (ODEs)], A.A. Andronov and L.S. Pontrjagin introduced, in 1937, the concept of *roughness* or *structural stability* (somewhat related to the *robustness* notion in control engineering). The importance of this concept is essential both in practice and in theory. An ODE (a dynamic system) is said structurally stable if the topological structure of its solutions does not change for small modifications of its parameters or of its structure. To be physically significant, a model of dynamic system must respect the following conditions:

1. A solution should exist
2. This solution should be unique
3. The unique solution should be continuous with respect to the data contained in the initial conditions or in the boundary conditions
4. The dynamic system should be structurally stable

The first three conditions were formulated by Hadamard in 1923. The study of the problem of structural stability (or roughness) can be considered complete for the two-dimensional autonomous ODEs. Andronov and Pontrjagin formulated, in 1937, the corresponding basic theorems in the analytic case. They are given in Andronov et al. [11], which also presents an exposition of the notion of *degree of structural instability*. In 1952, De Baggis presented proofs of these theorems in the more general case of smooth functions.

For autonomous two-dimensional ODEs (two-dimensional vector fields), general conditions of structural stability are:

1. The system has only a finite number of equilibrium points and limit cycles, which are not in a critical case in the Liapunov's sense (all the eigenvalues have real part different from zero).
2. No separatrix joins the same, or two distinct equilibrium saddle points (i.e., one eigenvalue is positive, the other negative). In this case it is possible to define, in the parameter space of the system, a set of cells inside each of which the same qualitative behavior is preserved.

The knowledge of such cells is of primary importance for the analysis and the synthesis of dynamic systems in physics or engineering. On the boundary of a cell, the dynamic system is structurally unstable, and for autonomous two-dimensional systems (two-dimensional vector fields),

structurally stable systems are dense in the function space. Till 1966, the conjecture of the extension of this result for higher dimensional systems was admitted. But Smale [154] showed that, in general, this conjecture is false. Therefore, it appears that, with an increase of the problem dimension, one also has an increase of complexity of the parameter (or function) space. The boundaries of the cells defined in the phase space (such cells are basins), as well as in the parameter space, in general have a complex structure, which may be *fractal* (self-similarity properties) for n -dimensional vector fields, $n > 2$. The Smale sufficient conditions of structural stability were the subject of new research in Russia, in particular with the Shilnikov's results (see the corresponding references in Shilnikov [148–150]).

The second approach of nonlinear dynamics corresponds to the *analytical methods*. Here, the aforementioned complex transcendental functions are defined to be convergent, or at least asymptotically convergent series expansions, or in “the mean.” The method of Poincaré's small parameter, the asymptotic methods of Krylov–Bogoliubov–Mitropolski are analytical. So are the averaging methods, and the method of harmonic linearization in the theory of nonlinear oscillations.

The two nonlinear dynamics approaches constitute relatively independent branches of the *nonlinear oscillations theory*. They have the same aims: construction of mathematical tools for the solution of concrete problems; and development of a general theory of dynamic systems. Since 1960, the important development of computers has provided a large extension to the numerical approach. Such an approach constitutes a powerful tool when associated with the qualitative and analytical methods.

During the last 30 years, interest in *deterministic models* generating solutions without any regular character (called *chaotic behaviors*, since 1975) has been increasing. It is about behaviors sensitive with respect to initial conditions, and very small parameter variations. Such a sensitivity induces a *practical unpredictability* of the model's behaviors solutions, due to the “physical” finite precision related to the data of a concrete problem. Most scientific fields dealing with dynamical processes have been plagued by such problems. More recently, it has been observed in the case of electronics, signal processing, and control. *Chaotic dynamics* is a subset of the interdisciplinary field named “*complex dynamics*,” a domain of the nonlinear dynamics concerned with the study of systems with nonlinearities inducing strong effects. Indeed, among dynamical behaviors met in the different scientific fields, one can discern those for which nonlinearities generate small effects and those for which they are dominant. In the latter case, the problem of the *transition order-chaos* under the influence of small parameter variations, or a small structure perturbation, gains particular importance. Generally, this gives rise to characteristic infinite sequences of *bifurcations*. Such bifurcation sequences are diverse and nonclassical. They result from an increase of nonlinear effects in the sense order toward chaos.

To avoid an abstract mathematical definition, a motion generated by a *purely deterministic system* will be chaotic if it does not represent any dynamical regularity and if it is sensitive with respect to very small changes of initial conditions (phase space). Such a behavior is also accompanied by a sensitivity with respect to very small parameter changes (parameter space), or to very small changes of a model structure (function space). For concrete systems, due to the “physical” finite precision of the data related to the phase and the parameter spaces, the corresponding models lose their *predictability capacity*, which is an essential characteristic of chaotic processes. In contrast to other fields where chaotic behaviors are accepted as a natural effect, generally in engineering they are considered as unfavorable, but in some applications they are used for obtaining special useful functions. In the two cases discussed earlier, a fundamental knowledge of bifurcation mechanisms generating chaos is essential for a good parameter choice at the synthesis step of an engineering project. This choice ensures that the chaos is either absent when it is unwanted or is present with fixed characteristics when it is related to a well-defined operating function. In both cases, it is taken into account the environment modifications.

The present chapter, essentially dealing with discrete models, is presented in the framework of the qualitative methods of nonlinear dynamics. Discrete models (*recurrences* or *maps*) are of two types with different properties. The first type corresponds to *invertible maps* T (i.e., the inverse map T^{-1} exists and is unique). The second type is related to *noninvertible maps* [i.e., depending on the phase (state) point either the inverse map T^{-1} is not real or it is not unique]. The latter situation induces new singularities and bifurcations. Note that a map (or a recurrence) model can either directly describe a system with discrete information or be the result of a *Poincaré section* (see Section 3.2) applied to an ODE. This text also deals with complex dynamics of continuous systems.

It is clear that a simple chapter, even in a survey form, cannot present a complete view of the title subject. The purpose of this text is limited to providing basic knowledge of chaotic phenomena with bifurcations generating such behaviors, in a non-abstract elementary form. In this framework, it is rather a guide to more complete information, which can be gathered from the references. This might provide additional theoretical and practical insight for researchers and engineers dealing with dynamical systems in different fields, among them control engineering, signal processing, and nonlinear electronics. Even if the volume of references here is relatively large, it does not pretend to be exhaustive. Complementary lists of references can be found in Gumowski and Mira [68, 69], Guckenheimer and Holmes [64], Abraham and Marsden [2], Mira [108, 111, 112], Sharkovskij et al. [146], Mira et al. [121], [130] Shilnikov et al. [150], Rössler [141], and Ueda [160–162].

For basic formation on nonlinear dynamics, Andronov et al. [10], using simple mathematics and useful applications for engineers, must be considered as the Bible. A reader without any knowledge about nonlinear dynamics and without any liking for mathematics can acquire great deal of information about this field, extended to chaotic behaviors, from Abraham and Shaw [3]. It is a pedagogical book, which uses only geometric figures and relations with concrete mechanical and electronic circuits, to describe the most complex nonlinear behaviors. Readers intrigued by more sophisticated mathematics can consult (among other books) Abraham and Marsden [2], Guckenheimer and Holmes [64], Sharkovskij et al. [146], and Shilnikov et al. [150]. Numerical algorithms for nonlinear dynamics are presented in Kawakami [79], Parker and Chua [135], Carcassès [30–32], Carcassès and Kawakami [33, 34]. A part of the results presented here are due to a group whose research was conducted in Toulouse from 1963 to 1996. The history of this group (now reduced to the author), with an extended presentation of the problems and references, is given pp. 95–197 of Abraham and Ueda [5]. This book is devoted to history of “teams and people who had struggled with chaos concepts before the acceptance of the new paradigm” (cf. the editors’ preface p. v and Abraham [1], Ueda [160–162], Rössler [141], Li and Yorke [90], and Smale [156]).

The remainder of the chapter is organized as follows. In Section 1.2, a very simple example (a one-dimensional quadratic recurrence or map) tackles the chaos problem and the related unpredictability thus generated. This is completed by a short description of associated behaviors in the general case. Section 1.3 deals with some generalities on discrete models. They can have different forms (explicit, implicit, parametric, autonomous, nonautonomous, invertible, and noninvertible) corresponding to a large variety of applications. References to some of them are provided. They can also be *indirect discrete models* associated with an ODE by a Poincaré section, which decreases the original dimension of the problem. Section 1.4 defines the singularities and the bifurcations common to invertible and noninvertible maps. When the map comes from an ODE, the sense of these notions is given in the continuous case. Singularities and bifurcations specific to noninvertible maps are discussed in Section 1.5. Section 1.6 presents two notions specific to noninvertible maps, having a practical interest: that of *absorbing domain* and of a *chaotic domain*. A survey on basin properties and their bifurcations is given in Section 1.7. Section 1.8 briefly presents map models having vanishing denominators. Such maps introduce new singularities and new bifurcation types. These results also concern maps T without vanishing denominators, but the inverse, T^{-1} , has a vanishing denominator, which gives rise to characteristic chaotic attractors. Section 1.9 tackles the difficulty in distinguishing a purely chaotic behavior (deterministic origin) from a noise effect. In the case of coexistence of these two phenomena, the extraction of the chaotic signal presents major difficulties. References to this problem and the characterization (different

definitions of the *dimension* of a strange attractor, *Liapunov exponent*) of chaotic behaviors are presented.

1.2 Chaos and Unpredictability

As noted in Section 1.1, the unpredictability property of chaotic behaviors is essentially related to a large sensitivity of model solutions with respect to initial conditions. The simplest example illustrating this point is the following discrete model under the form of a recurrence relationship: a noninvertible map called *Myrberg's map* [126–129]:

$$x_{n+1} = x_n^2 - \lambda, \quad \lambda = 2, \quad x(n=0) = x_0, \quad n = 1, 2, 3, \dots \quad (1.1)$$

In the interval $-2 < x < 2$, its solution remains bounded and chaotic. It is an exceptional case which can be stated from a classical transcendental function of the mathematical analysis

$$x_n = 2 \cos \left[2^n \arccos \left(\frac{x_0}{2} \right) \right]$$

where an appropriate determination of $\arccos$ is chosen for each initial condition x_0 . Here, the chaos is generated by the ordinates of a periodical function taken at exponentially increasing abscissa. The sensitivity of the solution x_n with respect to the initial condition x_n can be defined by the coefficient

$$S_n = \frac{\partial x_n}{\partial x_0} = -2^{n+1} \left(1 - \frac{x_0^2}{4} \right)^{-1/2} \sin \left[2^n \arccos \left(\frac{x_0}{2} \right) \right]$$

Therefore, S_n is a function of n , quickly increasing on the whole due to the term 2^{n+1} . Two very close initial conditions x_0, x'_0 of the interval $]-2; 2[$ generate two bounded iterated sequences, but $|x_n - x'_n|$ increases quickly when n increases for $n < N$. When $n > N$, this difference ceases to increase (because x_n is bounded) and varies in a muddled way. The segment $-2 \leq x \leq 2$ is called *chaotic segment*. It is characterized by the presence of many infinite sequences of unstable cycles (i.e., points such that $x_{n+k} = x_k$, $x_{n+p} \neq x_p$ for $p < k$) with increasing period k , and their limits when $k \rightarrow \infty$. The resulting point set has a *fractal* organization, that is the set is *self similar* (the whole is similar to the parts, even if they are infinitesimal). One-dimensional quadratic maps, obtained by a linear change of variable [as the so-called “logistic map” $x' = \lambda x(1 - x)$], have the same properties. The map (1.1) gives a one-dimensional example of

an intrinsically deterministic behavior with a chaotic dynamics, related to very large sensitivity with respect to very small variations of the initial state. This behavior clearly appears through such a simple example from the very rare possibility of having an analytical form of the solution, using the classical transcendental functions (see other cases pp. 24–31 of [108] and pp. 33–45 of [121]). More generally, chaos generated by one-dimensional maps are discussed in Li and Yorke [90] and references presented therein.

Generally, whatever be the model's nature, or the process (continuous or discrete) or its dimension, the solution is a nonclassical, nontabulated function of mathematical analysis. However, the existence of infinitely many sequences of unstable periodic stationary states is one of the characteristic features of chaotic behaviors with a deterministic origin. Such a chaotic behavior can be either stable (*strange attractor*) or unstable (*strange repeller*). A strange repeller leads either to a *chaotic transient* toward a nonchaotic stable stationary state or to *fractal basin boundaries* (*fuzzy boundaries*) separating the influence domain of m asymptotically stable stationary states. An initialization in a region of fuzzy basin boundary leads to uncertainty about the convergence of the model state toward one of the m stable stationary states, after a chaotic transient which persists as long as the state does not leave the fuzzy boundary region. The case of the chaotic transient toward only one stable stationary state is unexpected in the sense that a short term forecast is very difficult, but not a long-term one as the transient ends in a regular convergence toward this stable stationary state.

From a practical point of view, a stable periodic motion, having a period larger than the possible duration of observations with an irregular evolution during this period, can be also considered to be chaotic. In such a case (the simulation one) the term “chaos in a nonstrict sense,” or *nonstrict chaos* will be used with respect to the *strict chaos*, for which it is mathematically possible to prove the existence of chaos. When an erratic dynamical behavior is experimentally observed, the following fundamental question arises: is its origin deterministic, or coming from a purely random source, or mixed?

1.3 Generalities on Discrete Models

1.3.1 Different Forms of Models

This section concerns dynamical systems, described directly or indirectly (link with ODEs defined below) by a “discrete” equation, whose solution is a sequence of points determined by an initial point (initial condition). The model of many processes are of such a type called *recurrence relationship*

(or simply *recurrence*). The explicit form is written as:

$$X_{n+1} = F(X_n, \Lambda) \quad X(n=0) = X_0 \quad (1.2)$$

where X is a phase (or state) vector, Λ is a parameter vector, and X_0 is the initial condition. Depending on the scientific field, such an equation is also called *iteration*, or *map*, or *point-mapping*. Sometimes it is wrongly called *difference equation* because the solution of $X(t)$ is no longer a point sequence and is defined by an initial function $X(t) = X_0(t)$ for $-1 \leq t < 0$ (cf. Sharkovskij et al. [146]). In this chapter, relationship (1.2) will be called a map, denoted T . Its equation is symbolically represented by:

$$X_{n+1} = TX_n \quad \text{or without lower indices} \quad X' = TX \quad (1.3)$$

The point X' (or X_{n+1}) is called the *rank-one image* (or rank-one consequent) of X (or X_n). It is worth noting that the single-valuedness (X' exists and is unique) of the function $F(X_n, \Lambda)$, defining the map T , does not imply anything about the existence and uniqueness of its inverse $X = T^{-1}X'$. Indeed, this inverse may not exist, or it may be multivalued, then the map is called *noninvertible*. The map is *invertible* if its inverse exists and is unique. Considering the inverse map, $X = T^{-1}X'$ belongs to the set of *rank-one preimages* (or rank-one antecedent) of X' , which may be made up of several points, or only one point, or even void. The map $X_{n+r} = F_r(X_n, \Lambda)$ deduced from relationship (1.2) after r iterations is denoted T^r , $X_{n+r} = T^r X_n$, where X_{n+r} is the *rank- r image* (or rank- r consequent) of X_n . A point X_n belongs to the set of *rank- r preimages* (or rank- r antecedents) of X_{n+r} . The discrete time n does not appear explicitly in Equation (1.2) and Equation (1.3), so they are called *autonomous*. The equation is called *nonautonomous* if n appears explicitly:

$$X_{n+1} = F(X_n, n, \Lambda), \quad X(n=0) = X_0, \quad X' = T_n X$$

The one-dimensional example: $x' = x^2 - \lambda$ (Myrberg's map), where λ is a parameter, illustrates a case of noninvertible maps and the inverse map T^{-1} is given by $x = \pm\sqrt{x' + \lambda}$. The rank-one preimage of a point x' is double-valued for $x' > -\lambda$, and is not real for $x' < -\lambda$. The point $x' = -\lambda$ is called *critical point* (in the Julia–Fatou sense). In many publications, the extremum $x = 0$, point of two coincident rank-one preimages, is also called critical, which will not be the case in this chapter. When the inverse map T^{-1} is multivalued, or may not exist, the map T is called *noninvertible* or *endomorphie*. When T^{-1} exists and is unique, the map T is said *invertible*, and if the map is smooth, T is called a *diffeomorphism*.

More generally, a discrete model of a process in engineering does not have the aforementioned, simple explicit form. It presents itself either in

an implicit form or in a parametric form, frequently of a noninvertible type. The first case corresponds to the following relation:

$$F(X_{n+1}, X_n, \Lambda) = 0 \quad (1.4)$$

The parametric case may give rise to different formulations:

$$F(X_n, \Lambda) = G(V_n, \Lambda), \quad G(X_{n+1}, \Lambda) = F(V_n, \Lambda),$$

$$\dim X = \dim V \quad (1.5)$$

$$\Phi(X_n, X_{n+1}, V_n, \Lambda) = 0, \quad \dim X = m, \quad \dim V = s,$$

$$\dim \Phi = m + s \quad (1.6)$$

$$X_{n+1} = F(X_n, V_n, \Lambda), \quad G(X_n, V_n, \Lambda) = 0, \quad \text{with } \dim X = m,$$

$$\dim V = \dim G = s, \quad m \geq s \quad (1.7)$$

$$\Phi_i(X_n, X_{n+1}, V_n, \Lambda) = 0, \quad \dim X = m,$$

$$\dim V = s, \quad i = 1, 2, \dots, s + 1 \quad (1.8)$$

In Equation (1.5) to Equation (1.7), V is the “auxiliary parameter” of the parametric form, which has a different nature with respect to that of the “natural parameter” Λ . When they are numerically treated, these map equations do not give rise to more difficulties than the explicit form. In (1.7), V is frequently a discrete time (for $m = 1$, it is generally a commutation time) defined by $G(X_n, V_n, \Lambda) = 0$. The choice of one solution belonging to V defined by $G(X_n, V_n, \Lambda) = 0$ (if it is not unique) is assured from “physical” conditions associated with this relation. It is the same for the choice of X_{n+1} in (1.4)–(1.6) and (1.8). The solution of (1.2) to (1.8), for the initial condition $X(n = 0) = X_0$, is a sequence of points:

$$X_n = X(n, X_0, \Lambda), \quad n = 1, 2, 3, \dots \quad (1.9)$$

which is called *iterated sequence*, or *discrete phase trajectory*, or *orbit*. The map T can be considered to be an implicit definition of the function $X(n, X_0, \Lambda)$. Though theoretically quite satisfactory, such a definition is practically almost useless, because generally the function X is unknown, except for the linear case and for very few examples in the nonlinear case. In all non-contrived cases, it cannot be expressed explicitly in terms of the known elementary and transcendental functions.

Many classes of discrete dynamical systems give rise to models in the form of invertible or noninvertible maps in engineering, physics, computing and numerical simulation, dynamics of population, economics, biology, etc.

In control engineering, it is particularly the case of:

Systems using sampled data [35, 113], or switching elements [56], or pulse modulation (width modulation, frequency modulation) (cf. [68] and pp. 348–356, 366–370 of [109])

Adaptive control [6, 7, 48, 49, 85, 86]

Neural networks [39, 137, 139]

In nonlinear electronics, it is about rectifier using thyristors with voltage feedback, or current feedback [56], pp. 447–460 of [68], and pp. 370–387 of [109], oscillations [91–93], and Chua's circuit (Chua [40]).

Signal processing is concerned with bifurcations in the DPCM transmission system with an order two predictor [47, 55], sigma-delta modulation (described by piecewise continuous maps) [45], digital filters [133], and chaos synchronization [71, 134, 136, 151] for secure communications [99]. Frequently, in the aforementioned engineering fields, the function $G(X_n, V_n, \Lambda) = 0$ is the time interval separating the indices n and $n + 1$ [56] and pp. 366–387 of [109]. The function $G(X_n, V_n, \Lambda) = 0$ can also be an integral, one of whose bounds is V_n (case of the IPFM, integral pulse frequency modulation). In physics, “indirect” discrete models are for problems of turbulence, radiophysics, etc., via reduction of boundary value problems to difference equation [93, 145–147]. Economics and biology often lead to non-invertible maps [14, 52, 62].

A discrete equation of the earlier type corresponds to a direct model of a dynamic system or constitutes an indirect description of a continuous process. By *direct model* it is supposed that Equation (1.3) or Equation (1.2) is that of a dynamic system, which by its own nature is of discrete type, that is, the available information about its evolution is only accessible in a sampled form (discrete time). By *indirect description* it is assumed that the discrete equation is associated with an ODE, with the aim of obtaining an easier study of the original equation. This approach presents two different aspects. The first one concerns the discretization methods of ODEs, leading in particular to numerical simulations of continuous processes. In this case, depending on the method, the dimension of (1.3) is either equal to that of the ODE or higher. With the second aspect, the map is the result of application of the classical Poincaré's method of section surface to an ODE whose “real” dimension is m , this with the aim of a decrease of the initial real dimension. Then the map dimension is $m - 2$ if the ODE is conservative, and $m - 1$ if it is not.

1.3.2 Maps Obtained from an ODE by a Poincaré Section

Such maps correspond to an indirect description of a model in an ODE form after applying the Poincaré's section method. For the sake of simplicity, we

limit to a “real” ODE dimension equal to three. Therefore, this method can be understood by considering the following dynamical systems (where t is the time):

$$\frac{du_i}{dt} = f_i(u_1, u_2, u_3), \quad i = 1, 2, 3 \quad (1.10)$$

$$\frac{du_i}{dt} = g_i(u_1, u_2, \omega t), \quad i = 1, 2 \quad (1.11)$$

with f_i and g_i being smooth functions of their (real) arguments and g_i being periodic in t with the period $\tau = 2\pi/\omega$. Equation (1.10) is a three-dimensional autonomous ODE and (1.11) is a two-dimensional non-autonomous ODE. Nevertheless, they can be considered as having the same *real dimension* $m = 3$ because, by adding into (1.11) a third relation $du_3/dt = \omega$, an equivalent form of (1.10) is obtained.

Let U be the phase vector: $[u_1, u_2, u_3]$ of (1.10), and the phase vector $[u_1, u_2]$ for (1.11). Considering an initial condition $U = U_0$ for $t = t_0$, the phase (or state) trajectory is the curve of the phase (or state) space defined by the solution $U = U(t, U_0)$ of the earlier two ODEs.

With (1.10) considering the three-dimensional space (u_1, u_2, u_3) , a “regular” surface S transverse intersects the whole set of phase trajectories and a point M_n ($t = t_n$) intersects $U = U(t, U_0)$ with S . Let (x_n, y_n) be the M_n coordinates defined from the reference axes related to S . Let N be oriented normal to S . For increasing values of time t , let M_{n+1} ($t = t_{n+1}$) be the following intersection of $U = U(t, U_0)$ taking place in the same sense as the U evolution from M_n [i.e., the scalar products $NU(t_n, U_0)$ and $NU(t_{n+1}, U_0)$ have the same sign]. Then, the points M_n and M_{n+1} are related by an autonomous two-dimensional map

$$M_{n+1} = T M_n, \quad n = 0, 1, 2, \dots$$

obtained from the solution (generally a numerical one) in the interval (t_n, t_{n+1}) . For the ODE (1.11), the same form of map is defined considering the points M at times $t_n = n\tau$ and $t_{n+1} = (n+1)\tau$:

$$M_n [x_n = u_1(n\tau), y_n = u_2(n\tau)] \quad \text{and} \quad (1.12)$$

$$M_{n+1} \{x_n = u_1[(n+1)\tau], y_n = u_2[(n+1)\tau]\}$$

In the two cases, the study of ODE with a real dimension of three can be made from the associated two-dimensional autonomous map $M_{n+1} = T M_n$. This result is extended to all ODEs, whatever be their dimension. This leads to a decrease of one unit from the original real dimension of the ODE, which facilitates its study [72, 73, 79, 84]. Practically, in the

general case, the solution $U = U(t, U_0)$ is not analytically known. Nevertheless, this is not a difficulty because between the times t_n and t_{n+1} the ODE can be solved numerically using a computer, which numerically defines the map T on which all the operations of the following sections can be programmed. When the ODE data are smooth the map T is always invertible.

1.4 Singularities and Bifurcations Common to Invertible and Noninvertible Maps

1.4.1 Singularities and Bifurcations

Let T be a p -dimensional map (1.2) or (1.3) depending on the parameter vector Λ . Let X_0 be an initial condition. Qualitative methods of nonlinear dynamics are used to characterize the nonclassical transcendental function $X(n, X_0, \Lambda)$ of (1.9). A meaningful characterization consists of the identification of its *singularities* and the behavior of these singularities so-obtained, or any change of their qualitative properties, is called a *bifurcation*. In the parameter space, the boundary-separating behaviors of X_n , which are qualitatively different, are called a set of *bifurcation values* of the system parameters, for which the system is structurally unstable.

The simplest singularities are zero-dimensional: *period* (or *order*) k -cycles, denoted also k -cycles. A k -cycles is a set of k consecutive points X_i^* , $i = 1, 2, \dots, k$, permuting by successive applications of T , such that $X_i^* = T^k X_i^*$, with $X_i^* \neq T^r X_i^*$ for $1 \leq r < k$. When $k = 1$ the point X^* is called a *fixed point* (period-one cycle). In the rest of the chapter, when “cycle” is used, this may implicitly concern a fixed point. A cycle may be attracting (stable), or repulsive (unstable).

Let T be a smooth map. Then, it is possible to define the Jacobian matrix at a fixed point X^* , and considering T^k the Jacobian matrix at a period k -cycle point X_i^* . The p eigenvalues S_j , $j = 1, \dots, p$, of such a matrix are called the fixed points, or the cycle, *multipliers*. A cycle is stable, if and only if, all the multipliers are such that $|S_j| < 1$. It is unstable when at least one of the multipliers is $|S_l| > 1$. When at least one of the multipliers is $|S_l| = 1$ for a parameter value $\Lambda = \Lambda_b$, it corresponds to a critical case in the Liapunov' sense. Crossing through this case by a Λ variation gives rise to a *local bifurcation*. An unstable cycle with $|S_r| > 1$, $|S_s| < 1$, $\dim r + \dim s = p$, is called a *saddle*. The dimension s and the sign of each multiplier define different types of saddle. A fixed point, or a cycle, such that all the multipliers are as $|S_j| > 1$, $j = 1, \dots, p$, is said to be *expanding*.

When $\dim X = p = 2$, $X = (x, y)$, according to their multiplier values, cycles are classified into:

Stable (resp. unstable) node if the multipliers are real with $|S_1| < 1$ and $|S_2| < 1$ (resp. $|S_1| > 1$ and $|S_2| > 1$). A node is of *type one* if $S_1 > 0$ and $S_2 > 0$, of *type two* if S_1 and S_2 have opposite signs, and of *type three* if $S_1 < 0$ and $S_2 < 0$

Focus if the multipliers are not real

Saddle if $|S_1| < 1$ and $|S_2| > 1$. A saddle is of *type one* if $S_1 > 0$ and $S_2 > 0$, of *type two* if S_1 and S_2 have opposite signs, and of *type three* if $S_1 < 0$ and $S_2 < 0$

A period k -cycle is identified by the symbolism $(k; j)$, j being an index characterizing the permutation of the k -cycle points by k successive applications of T . This index permits the differentiation of cycles having the same period k and issued from different bifurcations (two cycles coming from the same bifurcation have the same permutation points). Invariant curves $G(x, y) = c$, c being a constant, by T (resp. T^k), passing through a fixed point (resp. period k -cycle) are manifolds, solutions of the functional equation $G(x_n, y_n) = G(x_{n+1}, y_{n+1})$ [resp. $G(x_n, y_n) = G(x_{n+k}, y_{n+k})$]. The complexity of this classification increases with the map dimension.

Manifolds (or sets) of dimension $d = 1, 2, \dots, p - 1$ ($\dim X = p$), invariant or mapped onto itself, by T or T^{-1} (resp. T^k or T^{-k}), and passing through a cycle point, constitute singularities of higher complexity with respect to fixed points and cycles. Locally, they are defined from the eigenvectors associated with the cycle multipliers (if they are real). For a saddle cycle X^* , the manifold (or set) associated with $|S_s| < 1$ is called the *stable manifold* (or *set*) $W^s(X^*)$ of this cycle. The manifold (or set) associated with $|S_r| > 1$ is called the *unstable manifold* (or *set*) $W^u(X^*)$ of the saddle cycle. Eigenvectors of a fixed point permit to define the local tangent manifold of such a point. Nonfractal singularities of dimension $p - 1$, invariant by T^{-1} , or T^{-k} bounding open regions of the p -dimensional phase (or state) space, inside each of them the qualitative behavior is well defined, play a fundamental role. These regions correspond to initial conditions giving rise to a transient toward a stable steady state. Generally, each of them constitutes the *influence domain* (called *basin*) $D(A)$ of a well-defined attracting set A . A closed and invariant set A is called an *attracting set* if some neighborhood U of A exists such that $T(U) \subset U$, and $T^n(X) \rightarrow A$ as $n \rightarrow \infty$, $\forall X \in U$. Generally, the basin boundary $\partial D(A)$ contains at least a saddle with $|S_s| < 1$, $s = 1, 2, \dots, p - 1$, and its stable manifold W^s . When the map T is invertible, a basin is always *simply connected*. This is not always the case when T is noninvertible, the basin being either *simply connected*, or *multiply connected*, or *nonconnected*.

A map may also generate singularities with a noninteger dimension. These singularities constitute what is called “*fractal sets*,” which can be attracting (*strange attractor*) or repulsive (*strange repeller*) for the points located in a sufficiently small neighborhood of such a set. Whatever be the map [invertible (with $p \geq 2$) or noninvertible (with $p \geq 1$)], a basin boundary $\partial D(A)$ can also be fractal (i.e., it can have a noninteger dimension). In this case, $\partial D(A)$ contains a strange repeller. The set $\partial D(A)$ is sometimes called *chaotic basin boundary*. If a map is noninvertible, a multiply connected or nonconnected basin generally implies a fractal basin boundary.

The set $W^s(X^*) \cap W^u(X^*)$ is called *homoclinic* if it is made up of an infinite number of intersections. Let X^* and Y^* be two fixed points (or cycles), then the set $W^s(X^*) \cap W^u(Y^*)$ is said *heteroclinic*. Homoclinic and heteroclinic situations are signs of (stable or unstable) chaotic behaviors. Bifurcations by *homoclinic* or *heteroclinic tangency* (limit of existence of infinite intersections) are *global bifurcations* which may correspond to bifurcations of an ordered dynamics toward a chaotic one.

Consider a m -dimensional dissipative system which, in the discrete case, is a diffeomorphism. An “ordinary” attractor A is a subset of the phase space so that in a sufficiently small neighborhood of A , an initial volume contracts and converges asymptotically toward A . In a *chaotic situation*, this contraction does not occur in all the directions. Indeed there is also a stretching toward certain directions, this leading to a complex folding of the initial volume, and giving rise to a foliated structure with infinitely many sheets, when the (continuous or discrete) time tends toward infinity. At the limits, a section in the direction of contraction locally gives a Cantor set. Then the final figure is fractal, and corresponds to a strange attractor with a noninteger dimension. In the two-dimensional case, this complex folding is related to the *Smale horseshoe* [152–156]. Such a process is also described in pp. 317–322 of [108] with application to the invertible quadratic map (diffeomorphism) $x' = 1 - ax^2 + y$, $y' = bx$.

1.4.2 Bifurcation Sets: Normal Forms of Exceptional Critical Cases

Consider the situation $\dim X = p = 2$, $X = [x, y]^t$, and a parameter plane $\Lambda(\lambda_1, \lambda_2)$. As mentioned earlier, the multipliers S_1 and S_2 of a $(k; j)$ -cycle are the eigenvalues of the linearization of T^k in one of the k points of this cycle.

In a parameter plane, a *fold bifurcation curve* $\Lambda_{(k)_0}^j$ is such that only one of the multipliers associated with a $(k; j)$ cycle is $S_1 = +1$. In the simplest case, this curve corresponds to the merging of a (k, j) saddle cycle ($S_1 < 1$, $S_2 > 1$) with a stable (or unstable) (k, j) node cycle ($0 < S_1 < 1$, $0 < S_2 < 1$). Similarly a *flip curve* Λ_k^j is such that one of the two multipliers is $S_1 = -1$, which gives rise to the classical period doubling from the $(k; j)$ cycle. In

the simplest case, this curve corresponds to a stable (k, j) node cycle ($-1 < S_1 < 0, S_2 < 1$) which turns into a (k, j) saddle k -cycle ($S_1 < -1, 0 < S_2 < 1$), giving rise to a stable $(2k, j')$ -node cycle ($0 < S_1 < 1, 0 < S_2 < 1$). Changing stable into unstable also results in a flip bifurcation.

As for a fold curve, a *Pitchfork bifurcation curve* $\Lambda_{(k)_0}^j$ corresponds to a $(k; j)$ -cycle with one of the multipliers $S_1 = +1$, but it is associated with three merging $(k; j)$ -cycles. For example, a stable $(k; j)$ -node cycle ($0 < S_1 < 1, 0 < S_2 < 1$) gives rise to a $(k; j)$ -saddle cycle ($S_1 < 1, S_2 > 1$) with two stable $(k; j)$ -node cycles ($0 < S_1 < 1, 0 < S_2 < 1$), with all these cycles merging for a parameter point on the pitchfork curve.

The case $S_i(X, \Lambda_b) = e^{\pm j\varphi}$, $i = 1$ or 2 , $j^2 = -1$, corresponds to a *Neimark bifurcation* (frequently and erroneously attributed to Hopf). In the simplest case, for example, when Λ crosses through Λ_b a stable (resp. unstable) focus point becomes unstable (resp. stable) and gives rise to a stable (resp. unstable) invariant closed curve (γ). The corresponding bifurcation curve (Γ_k^j) in the parameter plane is called a *Neimark curve*.

Fold, flip, pitchfork, and Neimark bifurcation curves are given in a parametric form [the vector X being the parameter of the parametric form, $S_i(X, \Lambda)$ being one of the two multipliers of the cycle (k, j) considered here] by the following relations:

$$X = T^k(X, \Lambda), \quad X \neq T^k(X, \Lambda), \quad \text{for } r < k, \quad \dim X = 2$$

$$S_i(X, \Lambda) = +1, \quad i = 1 \text{ or } 2, \text{ for fold and pitchfork curves}$$

$$S_i(X, \Lambda) = -1, \quad i = 1 \text{ or } 2, \text{ for flip curves}$$

$$S_i(X, \Lambda) = e^{\pm j\varphi}, \quad i = 1 \text{ or } 2, \quad j^2 = -1, \text{ for Neimark curves}$$

The Neimark bifurcation may give rise to several situations when φ is commensurable with 2π . The simplest one corresponds to the closed curve (γ) made up of an unstable (resp. stable) manifold of a period k saddle associated with a stable (resp. unstable) period k node (or a period k focus). More complex cases, depending on the nonlinear terms, occur when certain values of φ , commensurable with 2π , $\varphi = 2p\pi/q$, are related to *exceptional critical cases* requiring special *normal forms* for their study [12] (Holmes and Williams [75], pp. 215–239 of Mira [108], and Mira [101–103]). When the map is associated with an ODE, such cases may be related to complex resonance situations [118, 138]. More complex critical situations and their related bifurcations are described in pp. 239–255 of [108]. The aforementioned bifurcation curves correspond to *codimension-1 bifurcations*. Such curves may in turn contain singular points, the simplest ones being of codimension-2 (e.g., the *fold cusp* lying on a fold curve as meeting point of two fold arcs in a cusp form).

A set of bifurcation curves in a parameter plane $\Lambda(\lambda_1, \lambda_2)$ is not sufficient to account for the complete bifurcation properties. Indeed, it does not permit to identify the merging cycles. This is why the parameter plane must be considered to be made up of sheets. Each sheet is associated with a given cycle $(k; j)$ in a three-dimensional auxiliary qualitative space having a *foliated structure*. The third dimension is an adequate “qualitative” norm related to the (k, j) cycle. The identification of the sheet’s “geometry” consists of determining how it can be passed continuously from one sheet to another, following a continuous path of the parameter plane (i.e., to knowing the possible communications between sheets) (Mira et al. [114, 115, 117], Mira and Djellit [117], and Allam and Mira [8]). In the simplest case, a *fold bifurcation curve* is the junction of two sheets: one related to a saddle $(k; j)$ cycle, the other to a $(k; j)$ cycle having the modulus of each of the two multipliers less than 1 (stable node or stable focus) or having the modulus of its two multipliers greater than 1 (unstable node or unstable focus). A *flip bifurcation curve* is the junction of three sheets: one associated with a $(k; j)$ cycle having the modulus of its two multipliers less (resp. greater) than 1, the second sheet corresponding to a saddle $(k; j)$ cycle having one of its two multipliers less than -1 , the third being related to a $(2k; j')$ cycle having the modulus of its two multipliers less (resp. greater) than 1. A *pitchfork curve* is the junction of four sheets: three related to a stable $(k; j)$ node cycle ($0 < S_1 < 1, 0 < S_2 < 1$) and one related to a $(k; j)$ saddle cycle ($S_1 < 1, S_2 > 1$).

The sheets of the auxiliary three-dimensional space present folds along fold curves, and have junctions with branching along flip, or pitchfork, curves. The association of several bifurcation curves with their corresponding sheets and communications through codimension $s \geq 2$ singularities constitute a *bifurcation structure*. Codimension-2 points correspond to complex communications between the sheets. Therefore, the association of fold and flip curves in the neighborhood of a fold cusp leads to the definition of three fundamental communication types: the *crossroad area* (CRA), the *saddle area* (SAA), and the *spring area* (SPA) [108, 114]. Other types of singularities, with the corresponding three-dimensional representation of the sheets, are described in Carcassès et al. [29], Mira et al. [115, 119], Mira [110], Allam and Mira [8], and Mira and Qriouet [118]. However a representation of sheets of the parameter plane may lead to some difficulties due to the fact that generally the foliation is defined without ambiguities in a four-dimensional space. Thus, one may have situations such that it is possible to project the four-dimensional space into one of the three-dimensional spaces $(x, \lambda_1, \lambda_2)$ or $(y, \lambda_1, \lambda_2)$, and sometimes situations may exist for which this is impossible. In the latter case, three-dimensional projections may give rise to sheet intersections in the three-dimensional space which do not correspond to bifurcations.

The association of two of the aforementioned communication types leads to particular structures, or patterns, called *lip*, *quasi-lip*, *dovetail*, and *islands* [116, 117, 119, 123]. A lip ${}^m L_k^j$ results from the association of two fold segments, $\Lambda_{(k)_0}^j$ and $\Lambda_{(k)_0'}^{j'}$, of period k joining at two fold cusp points C_k^j and $C_k^{j'}$. These points with the related flip curves form a double crossroad area, or a double saddle area, or a double spring area, or an association between crossroad area–spring area.

Generally, the parameter vector Λ has a dimension higher than two. Therefore, varying the parameters different from (λ_1, λ_2) , the bifurcations organization in the parameter plane (λ_1, λ_2) undergoes qualitative changes. This means that transitions of an aforementioned structure into another are possible. So it was shown that a “crossroad area $\leftrightarrow$ spring area transition” may occur according to different mechanisms, identified from qualitative changes of a parameter plane (λ_1, λ_2) , and the associated three-dimensional foliated representation (Carcassès et al. [29], Mira and Carcassès [114], Mira [110], Allam and Mira [8], Mira and Qriouet [118], Mira et al. [119]). Useful algorithms permitting the determination of the nature of communication areas and their qualitative changes when a third parameter λ_3 varies, as well as the determination of different configurations of bifurcation curves and their foliated representation, are given in Carcassès [30–32] and Carcassès and Kawakami [33, 34], whatever be the map dimension. An algorithm for the determination of bifurcations by homoclinic or heteroclinic tangency can be found in Kawakami [78], Kawakami and Matsuo [81], and Yoshinaga et al. [165].

1.4.3 Singularities Sense when the Map is Obtained from a Poincaré Section

Consider a map T associated with an ODE such as (1.10) or (1.11), or with a real dimension larger than three, then:

A fixed point of T corresponds to either an equilibrium point of the ODE, or to a fundamental periodic solution, with a period τ for (1.11).

A period k -cycle of T corresponds to a *subharmonic oscillation* or to *fractional harmonic* (also called *ultra-subharmonic*) one, which is a periodic solution having a k -multiple period with respect to the earlier fundamental solution (see later for the definition of these two types of oscillations). In the case of (1.11) the period of the solution is $k\tau$.

A chaotic behavior of T corresponds to a chaotic behavior of the ODE solution [65, 73, 79, 159].

With ODE submitted to a periodic excitation (of external or parametric type), two cases must be considered. In the first one, the solution tends toward an equilibrium point in the absence of this excitation. Then, when the solution amplitude has a maximum, it is said that a *resonance* occurs. The second case corresponds to the presence of a stable periodic solution in the absence of periodic excitation. If this excitation exists, and if the ODE solution is periodic, we say that *synchronization* occurs. Subharmonic or fractional harmonic (ultra-subharmonic) oscillations can be either of resonance type or of synchronization type. It is worth noting that the knowledge of the period k of a cycle, of the index j (characterizing the permutation of its points by successive applications of T), and of its multipliers, does not provide complete information on the ODE periodic solution. It is necessary to associate with this information the knowledge of the solution during the period τ [i.e., the knowledge of the above (γ) closed curve].

Consider now a period τ solution of a fundamental solution, its Fourier series expansion, and the corresponding frequency power spectrum. Let r be the place occupied by a rank- m harmonic from an ordering based on the harmonics amplitudes in descending order. It is said that a rank- m resonance occurs when the amplitude of the rank- m harmonic occupies the place $r = 1$ in this ordering. *Higher harmonic oscillation* of rank- m [82, 83, 117] either of resonance type or of synchronization type, occurs when the place r of the rank- m harmonic is located at a position $r < m$, sufficiently far from m . Along a path of the parameter plane, the amplitude of harmonic lines (of the power frequency spectrum generated by the periodic solution), as well as their places r , varies continuously. It is possible to define curve arcs for which two harmonics of different ranks have the same amplitude with the place $r = 2$. The association of such arcs bounds regions of the parameter plane denoted *domains of (simple) predominance* of a rank- m harmonic. Inside each of these domains, such a harmonic has the place $r = 2$ in the ordering based on the amplitudes in descending order. When a point of the parameter space gives rise to a rank- m harmonic with the place $r = 1$, then it is said that this point belongs to a *domain of full predominance* of the rank- m harmonic [87, 88]. This situation corresponds to a *higher harmonic resonance*. A point of a domain of predominance, or full predominance, gives rise, in the continuous phase plane $[x(t), y(t)]$, to a closed curve (γ) , passing through the fixed point associated with the period τ solution. The higher the rank- m the more complex is the shape (γ) . Generally, the complexity is defined by (γ) self-intersecting loops: the more m increases the more the loop number increases. Such closed curves are given in Kawakami [79], Mira and Djellit [117], and Mira et al. [123].

A periodic solution of period $k\tau$ with $r = 1$ is called a $1/k$ -subharmonic [82, 83, 117]. A higher harmonic oscillation of rank- m [82, 83, 117] related to a period $1/k$ -subharmonic is called m/k -fractional harmonic. A fractional harmonic (or ultra-subharmonic) solution m/k is such that the dominant

frequency contained in $X(t)$ is $m\omega/k$, ω being the angular frequency of the aforementioned fundamental solution. This solution corresponds to a period k -cycle of T , but in the continuous phase plane $[x(t), y(t)]$ it gives rise to a closed curve (γ) passing through the k points of the cycle: the higher the m the more complex is the shape (γ). Such closed curves are discussed in Kawakami [79], Mira and Djellit [117], and Mira et al. [123]. Fractional harmonics are distinguished as *nonreducible fractional harmonics* (the ratio m/k cannot be reduced) and *reducible* ones. In the case of reducible harmonics, the ratio m/k can be reduced, but due to its relation with a k -cycle, it keeps this form to correctly identify its relation with a period k -cycle. Reducible harmonics have a more complex behavior, giving rise to specific bifurcation structures in a parameter plane [138, 158].

In the parameter plane, rank- m resonances, or synchronizations, $m = 1, 2, 3, \dots$, are related to the existence of an *isoordinal cascade* either of fold cusps or lips denoted ${}^m L_k^j$, $m = 1, 2, 3, 4, \dots$. “Isoordinal” means that each lip is made up of fold arcs corresponding to the same period k , ${}^m \Lambda_{(k)_0}^j$ and ${}^m \Lambda_{(k)_0}^{j'}$ joining at two cusp points ${}^m C_k^j$ and ${}^m C_k^{j'}$. Each fold cusp, or lip ${}^m L_k^j$, lies inside a rank- m domain of simple predominance. Such a cascade has a limit set corresponding to a rank- $m = \infty$ of the higher harmonic resonance [88, 117].

1.5 Map Singularities and Bifurcations Specific to Noninvertible Maps

1.5.1 Singularities and Bifurcations Induced by Noninvertible Maps

With respect to invertible maps, noninvertible maps T introduce a singularity of a different nature: the *critical set*. The rank-one critical set CM is the geometrical locus of points X having at least two coincident rank-one preimages. Such preimages are located on a set CM_{-1} , the set of merging (or coincident) of rank-one preimages. The set CM satisfies the relations $T^{-1}(CM) \supseteq CM_{-1}$ and $T(CM_{-1}) = CM$. A rank- q critical set CM_{q-1} is given by the rank- q image $CM_{q-1} = T^q(CM)$, $CM_0 \equiv CM$. If $\dim X = p = 1$, CM is a rank-one *critical point* C . If $\dim X = p = 2$, CM is a rank-one *critical curve* LC .

Such new singularities play a fundamental role in the attractors and basins structure and in their bifurcations. It is the case of “contact bifurcations,” resulting from the meeting of two singularities of different nature: an invariant manifold (or set) by T or T^{-1} with a critical set. This situation generally gives rise to *global bifurcations*, which may be related to homoclinic and heteroclinic bifurcations [51, 53, 98, 121]. Most of the results

obtained till now concern the general class of maps of the plane $T: \mathbf{R}^2 \rightarrow \mathbf{R}^2$. For such maps, the critical set CM generally becomes a critical curve LC . In exceptional cases the critical curve may include isolated points. This is the case when the inverse map T^{-1} has a vanishing denominator [16–26] (see also Section 1.8).

The singularity critical curve constitutes the fundamental tool for the study of two-dimensional noninvertible maps. In general, LC is made up of several branches separating the plane into regions whose points have different numbers of rank-one preimages (or antecedents). Therefore, the plane $\mathbf{R}^2$ can be subdivided into open regions Z_i ($\mathbf{R}^2 = \cup_i \overline{Z_i}$, $\overline{Z_i}$ being the closure of Z_i), each point of Z_i having i distinct rank-one preimages. There is a class of maps such that a region Z_0 exists. The boundaries of the regions Z_i are branches of the rank-one critical curve LC , locus of points such that at least two determinations of the inverse map are merging. The locus of these “coincident first rank preimages” is a curve LC_{-1} , called *rank-one curve of merging preimages*. As in any neighborhood of a point of LC , there are points for which at least two distinct inverses are defined; LC_{-1} is a set of points for which the Jacobian determinant of a smooth map T vanishes. If the map is nonsmooth, LC_{-1} belongs to the set for which the noninvertible map T is not smooth. The curve LC satisfies the relations $T^{-1}(LC) \supseteq LC_{-1}$ and $T(LC_{-1}) = LC$.

The simplest case is that of maps in which LC (made up of only one branch) separates the plane into two open regions Z_0 and Z_2 . A point X belonging to Z_2 has two distinct preimages (or antecedents) of rank-one, and a point X of Z_0 has no real preimage. The corresponding maps are said to be of $(Z_0 - Z_2)$ type. In more complex cases a classification of noninvertible maps from the structure of the set of Z_i regions can be made [121, 122]. It is worth noting that the bifurcations organization, for example in a parameter plane, may be very different from that given by invertible maps. Indeed with respect to the invertible case, the noninvertibility also adds new bifurcation structures in a parameter plane due to the presence of new co-dimension 2 points related to cycles with real multipliers $S_1 = +1$ and $S_2 = -1$ [38].

1.5.2 Fractal Bifurcations Structure of “Embedded Boxes” Type and Chaotic Behaviors

As discussed earlier, chaotic dynamics can be met in a discrete model having the lowest dimension p (i.e., $p = 1$). The necessary condition of such a behavior is that the corresponding map be noninvertible. Chaotic solutions appear with an invertible map only if the map dimension is at least equal to two. So at equal dimension, noninvertible maps present intrinsically better conditions favoring the birth of chaos [142, 144]. As shown

in Section 2, for one-dimensional quadratic maps, the chaotic behavior in an interval is due to the presence of infinitely many infinite sequences of unstable cycles with increasing period k , and their limit set (of “class 1”) being when $k \rightarrow \infty$. The infinite set of points of class 1 in its turn has infinitely many limit sets of class 2. Equivalently, limit sets of class $q \rightarrow \infty$ are defined [108]. This leads to a fractal organization of the whole set of repulsive singularities; that is, the set is *self similar* (the whole is similar to the parts even if they are infinitesimal). The Myrberg’s map (1.1) $x' = x^2 - \lambda$ illustrates the bifurcation sequences leading to such a situation.

The “classical” singularities of the solution of the one-dimensional quadratic map (1.1) are constituted by two fixed points (real if $\lambda \geq -1/4$) verifying $x' = Tx$, the cycles points of period (or order) k , $k = 2, 3, 4, \dots$, and their limit (fractal) sets when $k \rightarrow \infty$. The “nonclassical” singularities, which play an essential role in the fractal bifurcation structure, and the global bifurcations generated by this map, are made up of the set of the critical points of rank C_r , $r = 1, 2, 3, \dots$. With (1.1) Myrberg has been the first to show a series of essential results for the theory of dynamic systems:

- All the bifurcations values of (1.1) occur into the interval $-1/4 \leq \lambda \leq 2$.
- The number N_k of all possible cycles having the same period k , and the number $N_\lambda(k)$ of bifurcation values giving rise to these cycles, increases very rapidly with k . So one has $N_k = N_\lambda(k) = 1$, if $k = 2$; $N_k = 2, N_\lambda(k) = 1$, if $k = 3$; $N_k = 3, N_\lambda(k) = 2$, if $k = 4$; $N_k = 6, N_\lambda(k) = 5$, if $k = 5$; $N_k = 99, N_\lambda(k) = 28$, if $k = 10$; $N_k = 35,790,267, N_\lambda(k) = 7,895,679$, if $k = 30$; $N_k \rightarrow \infty, N_\lambda(k) \rightarrow \infty$, if $k \rightarrow \infty$ (for more details see Mira [108]).
- The cycles $(k; j)$ with the same period k differ from one another by the cyclic transfer (shift defined by the index j) of one of their points by k successive iterations by T . These cyclic shifts were defined by Myrberg using a binary code constituted by a sequence of $(k - 2)$ signs $[+, -]$ (*binary rotation sequence*). More or less explicitly, the Myrberg’s papers provide an extension of this notion to the case $k \rightarrow \infty$ and to general orbits (iterated sequences).
- For $\lambda < \lambda_{(1)s} \simeq 1.40115589, \dots$, the number of singularities is finite (T is said to be “Morse–Smale”). For $\lambda \geq \lambda_{(1)s}$, the number of singularities is infinite, and the situation is chaotic (stable or unstable chaos). The parameter $\lambda_{(1)s}$ is an accumulation value of bifurcations by period doubling (*Myrberg cascade* called “spectrum” by Myrberg [129] and often wrongly named *Feigenbaum cascade* [46]).
- The following cascades of bifurcations: “stable $(k2^i; j)$ -cycle $\rightarrow$ unstable $(k2^i; j)$ -cycle + stable $(k2^{i+1}; j')$ -cycle”, $i = 1, 2, 3, \dots$; k having a fixed given value; $k = 1, 3, 4, \dots$, occurs when λ increases.

- For $i \rightarrow \infty$, the bifurcation values $\lambda_b(k2^i; j)$, from a given period k , have a limit point $\lambda_{(k)s}^j, \lambda_{(1)s} < \lambda_{(k)s}^j < 2$, accumulation value by successive period doubling from a period k -cycle, $\lim_{i \rightarrow \infty} \lambda_b(k2^i; j) = \lambda_{(k)s}^j$.
- It is possible to classify all the binary rotation sequences via an ordering law (Myrberg's ordering law).
- A binary rotation sequence can be associated with the λ -value resulting from accumulation of bifurcations such that $i \rightarrow \infty$ or $k \rightarrow \infty$. This rotation sequence satisfies the ordering law.

All these fundamental results have been overlooked, in the contemporary papers dealing with this subject, which has created a very large void since 1978. Most of these results are now often attributed to the authors who rediscovered them later using another forms of quadratic maps, such as the logistic map or maps of the unit interval. Therefore, the characterization of a cycle or an orbit by a binary code was rediscovered by Metropolis [97] where the symbols “ R, L ” are introduced instead of Myrberg's symbols “ $+, -$ ”. It is also the case of the popular notions of *invariant coordinate*, *kneading invariant* related to properties of Myrberg's rotation sequences, now attributed to Milnor and Thurston [100].

The fractal “box-within-a-box” (or *embedded boxes*) bifurcation structure (*structure de bifurcation boîtes emboîtées* in French, see Mira [104, 108], Gumowski and Mira [68, 69], Mira et al. [121]) [63], generated by the map (1.1), corresponds to an ordering of the Myrberg cascades (or spectra). It is made from the nonclassical bifurcation $\lambda = \lambda_k^{j*}$ resulting from the merging of critical points $C_m, m = k, k + 1, \dots, 2k - 1$, with the points of a $(k; j)$ -cycle, which defines a limit of the above accumulation value $\lambda_{(k)s}^j$ when $k \rightarrow \infty$. Such embedded boxes bifurcation structures are also met for p -dimensional invertible or noninvertible maps, $p > 1$.

1.5.3 Homoclinic and Heteroclinic Situations: Their Bifurcations

Consider a p -dimensional noninvertible map T . Let U be a neighborhood of an unstable fixed point (saddle, node, or focus) p^* . The *local* (i.e., in U) *unstable set* $W_{\text{loc}}^u(p^*)$ of p^* is defined as the locus of points in U having a sequence of increasing rank preimages in U which tends toward p^* . The *global unstable set* is the locus of all the points for which a sequence of preimages exists, and converges toward p^* . It can be obtained by constructing the images of the local unstable set. When the map T is continuous and noninvertible with $p > 1$, the invariant unstable set $W^u(p^*)$ of a saddle point p^* is connected, but self intersections may occur (therefore, it may not be a manifold), which cannot happen for invertible maps. When $p = 2$, self

intersections and loops of $W^u(p^*)$ are as described earlier (pp. 373–374 of [68], pp. 203–222 of [69], and pp. 506–515 of [121]). The role of critical sets CM_q and sets CM_{-1} of merging preimages is again essential in understanding the formation of self intersections of the unstable set of a saddle fixed point and properties of invariant closed curves. Moreover, the bifurcation of an invariant closed curve turning into a chaotic attractor, by creation of local loops, is possible [121].

The stable set $W^s(p^*)$ of a saddle p^* is backward invariant $T^{-1}[W^s(p^*)] = W^s(p^*)$. It is mapped into itself by T , $T[W^s(p^*)] \subseteq W^s(p^*)$. It is invariant if T is invertible, while for a noninvertible map it may be strictly mapped into itself. When T is continuous, self intersections of $W^s(p^*)$ cannot occur (therefore, it may be called manifold, being either a connected manifold or the union of disjoint connected components). When T is noninvertible with $p = 2$, $W^s(p^*)$ may be nonconnected and made up of infinitely many closed curves passing through the increasing rank preimages of p^* . An equivalent property holds for higher dimensions $p > 2$.

A fixed point or a cycle is called *expanding* if all its multipliers are $|S_i| > 1$, $i = 1, \dots, p$, and if there exists a neighborhood U such that the absolute values of the jacobian matrix of T , or T^k , is larger than one for each point $X \in U$. In contrast to invertible maps the stable set W^s of an expanding point p^* can be defined [95]. It is made up of the arborescent sequence of increasing rank preimages of this point $W^s(p^*) = \bigcup_{n>0} T^{-n}(p^*)$. When a chaotic attractor exists, the unstable set W^u of an expanding fixed point is a domain [if an attractor exists W^u lies inside a *chaotic area* when $p = 2$ (see what follows)] bounded by pieces of critical sets CM_q , $q = 1, 2, \dots, r$.

A point q is said to be homoclinic to the non-attracting fixed point p^* (or homoclinic point of p^*) iff $q \in W^s(p^*) \cap W^u(p^*)$. Heteroclinic points are obtained when the stable and unstable sets are related to two different fixed points. As indicated earlier a “contact bifurcation” may correspond to homoclinic and heteroclinic bifurcations, and critical sets CM_q are useful for interpreting such problems. Classically, for invertible maps homoclinic and heteroclinic situations are defined for n -dimensional diffeomorphisms, $n > 1$, and only from saddle points. It is worth noting that the first “extended” notion (with respect to the classical one) of homoclinic and heteroclinic points in one-dimensional noninvertible maps, with an indication of its generalization for p -dimensional maps, $p > 1$, was introduced in Sharkovskij [143]. This was done by defining the stable set of a fixed point as the set of all its preimages of increasing rank. For a one-dimensional noninvertible map, the stable set of an unstable fixed point is made up of the infinite arborescent set of its preimages of increasing rank. The unstable set has at least one branch bounded by this fixed point and a critical point C_q . On the basis of these results, bifurcations by “homoclinic and heteroclinic contact” have been presented for the one-dimensional case (pp. 395–400 of [68] and pp. 294–296 of [108]) with, in an embryonic form, equivalences

of situations for higher dimensions. More extended references are given in Mira [112]. Thus homoclinic and heteroclinic sets, $W^s \cap W^u$, exist not only for saddle points but also for expanding points. More details on definitions and properties are given in Gardini [53] and pp. 13–21 of [121].

Homoclinic and heteroclinic sets are accumulation points of unstable cycles, when their period tend toward infinity, this leading to chaotic situations. In the case of ODEs, such situations result from an accumulation of infinitely many unstable periodic solutions of increasing period (i.e., from unstable subharmonics) and fractional harmonics whose ranks tend toward infinity. According to the case, the accumulation of unstable cycles, or of unstable periodic solutions for ODEs, gives rise [121, 122] either to: an attracting set, called strange attractor (case of the stable chaos) or to a repulsive set, called strange repeller (case of the unstable chaos). In the latter case two situations are possible: either that of a *chaotic transient* toward an attractor or that of a *chaotic basin boundary* (or *fuzzy boundary*) separating the basins of several basins. Such fractal sets have the specificity that their dimension is not an integer and are made up of an inextricable tangle of invariant sets related to unstable cycles with increasing period.

1.6 Absorbing Areas, Chaotic Areas, Bifurcations

1.6.1 Definitions and Properties

Consider a two-dimensional noninvertible map T . Critical curves permit to define the essential notions of *absorbing area* and *chaotic area* [37, 66–69, 80, 107, 108, 121]. Roughly speaking, an absorbing area (d') is a region bounded by critical curves arcs of finite, or infinite, rank LC_n , $n = 0, 1, 2, \dots, l$, $LC_0 \equiv LC$, such that the successive images of all points of a neighborhood $U(d')$, from a finite number of iterations, enter into (d') and cannot get away after entering. Except for some bifurcation cases, a chaotic area (d) is an invariant absorbing area whose points give rise to iterated sequences (or orbits) having the property of sensitivity to initial conditions. In general it contains infinitely many unstable cycles of increasing period, their corresponding limit sets, and the preimages of increasing rank of all these points. Its boundary ∂d is made up of LC_n arcs. Note that a chaotic area may be periodic of period k (i.e., constituted by k nonconnected chaotic areas invariant by T^k).

The role of critical curves is also fundamental in the definition of bifurcations leading either to the destruction or to a sudden and qualitative modification of absorbing areas and chaotic areas. In particular, such modifications concern: transitions “simply connected chaotic area $\rightarrow$ doubly

connected chaotic area" (or "annular area"), "nonconnected chaotic area $\rightarrow$ doubly connected chaotic area." A chaotic area (d) is destroyed via a bifurcation resulting from the contact of its boundary ∂d with the boundary $\partial D(d)$ of the basin $D(d)$ of (d). Then, as soon as it is destroyed, (d) turns into a strange repeller. All the bifurcation properties of such areas are presented in the aforementioned references. In numerical simulations, the LC_n arcs inside a chaotic area appears as a place of higher concentration of iterated points if the map is smooth or as separation of regions with different densities of iterated points if the map is not smooth. This characteristic is directly related to properties of local extremums of the map.

An extended notion of absorbing area and chaotic area, that of mixed absorbing area, mixed chaotic area, was also introduced in Barugola et al. [13] and Mira et al. [121]. These areas differ from the nonmixed ones by the fact that their boundaries are now made up of the union of critical curves segments and segments of the unstable set of a saddle fixed point, or a saddle cycle or even segments of several saddle unstable sets associated with different cycles. With respect to a "simple" (nonmixed) absorbing, or chaotic area, these areas are such that successive images of almost all points of a neighborhood enter into the area from a finite number of iterations and cannot get away after entering. The successive images of the points which do not enter into the area are those of the arc [out of (d)] of the stable set of the saddle point on the area boundary. Though not entering the area, these images tend toward the boundary saddle point.

Critical curves also play an essential role in the comprehension of the possibility of obtaining points of a same cycle located on both sides of an invariant closed curve (γ) (Frouzakis et al. [50], and pp. 534–537 of Mira et al. [121]). It is a "pathological" dynamical behavior, not encountered in invertible maps. Moreover, from an invariant closed curve (γ) infinitely many bifurcations (pp. 534–588 of [121]) give rise first to a *weakly chaotic ring* and later to a doubly connected chaotic area. Without an important enlargement, a weakly chaotic ring appears numerically as an invariant closed curve, but a section of the enlargement permits to discern a Cantor set.

1.6.2 Chaotic Areas: Microscopic and Macroscopic Points of View

Regarding chaotic areas, or mixed chaotic areas, it is important to emphasize that the purpose of the study of such areas is to obtain the "macroscopic" properties of the chaotic attracting set (defined earlier) leading to the considered area. In particular, these properties are those appearing in a first step from a numerical simulation of the iterated sequences generated by the map. The "microscopic" properties (i.e., the nature of closed invariant sets generated by such maps) or the internal structure of an attractor (if it exists), implies further studies and are more difficult to identify.

This concerns the set of *nonwandering points*: limit set of the unstable cycles with increasing period, limit set of their unstable set, and limit set of their preimages with increasing rank.

Considering the microscopic point of view, it worth noting that in 1979 a very important theorem was formulated [131, 132]. It states that in any neighborhood of a C^r -smooth ($r \geq 2$) dynamical system, in the space of dynamical systems (or a parameter space), there exist regions for which systems with homoclinic tangencies (then with structurally unstable or nonrough homoclinic orbits) are dense. Domains having this property are called *Newhouse regions*. This result is completed in Grochenko et al. [57] which asserts that systems with infinitely many homoclinic orbits of any order of tangency, and with infinitely many arbitrarily degenerate periodic orbits, are dense in the Newhouse regions of the space of dynamical systems. Such a situation has the following important consequence: systems belonging to a Newhouse region are such that a complete study of their dynamics and bifurcations is impossible. Indeed, in many smooth cases, due to the finite time of a simulation, what appears numerically as a chaotic attractor contains a “large” hyperbolic subset in the presence of a finite or an infinite number of stable periodic solutions. Generally, such stable solutions have large periods, and narrow “oscillating” tangled basins, which are impossible to exhibit numerically due to the finite time of observation, and unavoidable numerical errors. Thus it is only possible to consider some of the characteristic properties of the system, their interest depending on the nature of the problem [149]. Such complex behaviors occur for p -dimensional flows (autonomous ODEs) with $p > 2$, and thus for $p \geq 2$ invertible and noninvertible maps.

From a macroscopic point of view, the union of the numerous and even infinitely many stable solutions, which are stable cycles for a map, forms an attracting set denoted A . A numerical simulation of the map solution, by definition, is made from a limited number of iterations. Consider the case of a noninvertible map giving rise to a chaotic area, and the elimination of a transient, that is, the simulation is made after N iterations, N being sufficiently large to attain what at first glance appears to be a steady state. Then either the numerical simulation reproduces points of the chaotic area, related to a “strict” strange attractor in the mathematical sense, or it represents a transient toward an attracting set A including stable cycles of large period, a large part of them with a period larger than the simulation duration. The first case, for example, is that of some piecewise smooth maps (i.e., with isolated points of nonsmoothness), not permitting stable cycles (i.e., the Jacobian determinant cannot be sufficiently small). Assuming numerical iterations without error, in the second case the transient would be that toward a stable cycle having a period larger than the number of iterations, this transient occurring inside a very narrow basin, tangled with similar basins of the other stable cycles of large periods. In

the presence of unavoidable numerical errors, the iterate points cannot remain inside the same narrow basin. They sweep across the narrow tangled basins of the other cycles of the attracting set A . Then they reproduce a chaotic area bounded by segments of critical curves LC_q . This means that this chaotic area coincides with the numerical observation: in the *smooth case* as a transient toward the attracting set A located inside the area; in the *nonsmooth case* as a true (in the mathematical sense) strange attractor. Such a property constitutes an important characteristic of the system dynamics. This shows the high interest of the notion of chaotic area, even if in the smooth case it is impossible to numerically discriminate a situation of a strange attractor, in the mathematical sense, from that of an attracting set made up of stable cycles with very large periods.

1.7 Results on Basins and their Bifurcations

Let D be a basin, that is, the open set of points X whose forward trajectories (set of increasing rank images of X) converge toward an attracting set A . This notion, related to global properties of the map, is particularly important for applications. Considering a noninvertible map T , D is invariant under the backward iteration T^{-1} of T , but not necessarily invariant by T . The basin D and its boundary ∂D satisfy the relations:

$$T^{-1}(D) = D, \quad T(D) \subseteq D, \quad T^{-1}(\partial D) = \partial D, \quad T(\partial D) \subseteq \partial D$$

Here, the strict inclusion holds iff D contains points of a Z_0 region (i.e., with no real preimage). Such a basin may be *simply connected* as in the invertible case, but also *nonconnected*, and *multiply connected* [36, 120]. Its boundary ∂D may contain repulsive sets related to the presence of strange repellers SR. Such an unstable set SR is made up of infinitely many unstable cycles with increasing period, their limit sets of increasing class, the preimages of increasing rank of all these points. As indicated, a set SR gives rise to fractal basin boundaries (or fuzzy boundaries) separating the domain of influence of different attractors and chaotic transients toward a defined attractor [58–60, 105, 106, 121].

Since 1969, several papers have developed the role of critical curves in the bifurcations of type “simply connected basin $\leftrightarrow$ non-connected basin” (see pp. 228–261 of [68] and pp. 87–89 of [69]). It is the same for bifurcations of type “simply connected basin $\leftrightarrow$ multiply connected basin”. These basic bifurcations always result from the contact of a basin boundary with a critical curve segment and are generated by the same basic mechanism but in many different ways. They are equivalent to the simplest bifurcation, met in one-dimensional maps, with the merging of a

critical point with a point of a basin boundary. All these bifurcations and new ones, with detailed references (see also Mira [112]), as those related to the fractalization of basin boundary, are presented in Mira et al. [121].

1.8 Map Models with a Vanishing Denominator

To simplify the exposition, the map (of invertible or noninvertible type) is a two-dimensional one, and it is assumed that only one of the two functions defining the map T has a denominator which can vanish

$$T : x' = F(x, y), \quad y' = G(x, y) = \frac{N(x, y)}{D(x, y)}$$

where x and y are real variables, $F(x, y)$, $N(x, y)$ and $D(x, y)$ are continuously differentiable functions defined in the whole plane $\mathbb{R}^2$. Hence, the *set of nondefinition* of the map T (which is given by the set of points where at least one denominator vanishes) reduces to $\delta_s = \{(x, y) \in \mathbb{R}^2 | D(x, y) = 0\}$. It is assumed that δ_s is given by the union of smooth curves of the plane. The two-dimensional recurrence obtained by the successive iterations of T is well defined provided that the initial condition belongs to the set E given by $E = \mathbb{R}^2 \setminus \bigcup_{k=0}^{\infty} T^{-k}(\delta_s)$, where $T^{-k}(\delta_s)$ denotes the set of the rank- k preimages of δ_s [i.e., the set of points which are mapped into δ_s after k applications of T ($T^0(\delta_s) \equiv \delta_s$)]. Indeed, the points of δ_s , as well as all their preimages of any rank constituting a set of zero Lebesgue measure, must be excluded from the set of initial conditions that generate noninterrupted sequences by the iteration of the map T , so that $T : E \rightarrow E$.

Such a characteristic is the source of some particular dynamical behaviors, related to the presence of new kinds of singularities and bifurcations, as recently evidenced in Bischi et al. [19], where in particular the situation arising when $F(x, y)$ or $G(x, y)$ assumes the form $0/0$ in some points of $\mathbb{R}^2$ has been analyzed. In these references new singularities, called *focal point* and *prefocal curve*, have been defined which permit the characterization of specific geometric and dynamic properties, together with some new bifurcations. Roughly speaking, a *prefocal curve* is a set of points which are mapped (or “focalized,” as we shall say for short) into a single point, called *focal point*, by the inverse of T (if the map is invertible) or by at least one of the inverses (if the map is noninvertible). These singularities may also be important in the study of maps T defined in the whole plane (then without a vanishing denominator), but such that a determination T_i^{-1} of the inverse $T^{-1} = \bigcup_{i=1}^n T_i^{-1}$ has a vanishing denominator and possesses a focal point.

In particular, this situation gives rise to special kind of chaotic attractors, those presenting *knots* singular points (see Bischi et al. [19]; Figure 37).

Global bifurcations, due to the presence of focal points, cause the creation of structures of basins, specific to maps with a vanishing denominator, called *lobes* and *crescents*. They have been explained in terms of contacts between basin boundaries and prefocal curves [18, 19, 24, 26]. These structures have been recently observed in discrete dynamical systems of the plane arising in different contexts [20, 24, 27, 28, 54, 164].

The literature on chaotic dynamical systems mainly concerns bounded attracting sets, while unbounded trajectories are usually considered to be synonymous of diverging trajectories. Also, the definitions of attractor given in the current literature are almost all referring to compact sets [77, 140, 163]. The fact that this may be a restrictive point of view has been recently emphasized by some authors. For example, Brown and Chua [28] write "... in defining chaos, no restrictions as to boundedness is reasonable". Indeed, *unbounded chaotic trajectories* naturally arise in the iteration of maps with a denominator which can vanish. For example, the existence of a "nonbounded chaotic solution" in a one-dimensional recurrence with denominator has been shown in Mira [108] (see also p. 38 of [121]). The paper Bischi et al. [23] shows examples of unbounded chaotic trajectories and describe some nonclassical (or contact) bifurcations which cause the transition from bounded asymptotic dynamics to unbounded (but not diverging) dynamics, both in one-dimensional and two-dimensional fractional maps. The basic feature of an unbounded and not diverging trajectory is that points of arbitrarily large norm may belong to the trajectory, but they do not give rise to divergence (i.e., these points have images of smaller norm). Of course, this property may cause some difficulties in the numerical iteration of a map by a computer, since an overflow error may occur even if the numerically generated trajectory is not diverging. Furthermore, the occurrence of such a numerical error may be strongly dependent on the kind of computer or the kind of floating-point arithmetic used to perform the calculations. For this reason, even if the paper by Bischi et al. [23] gives some numerical representations of unbounded sets of attraction in order to help the reader to visualize the objects studied, the existence of unbounded chaotic trajectories is shown on the basis of theoretical arguments.

The study of peculiar dynamical behaviors of maps with denominator has been motivated by practical reasons, because discrete dynamical systems, obtained by the iteration of maps with denominator, occur often in applications. For example, many iterative methods for finding numerical solutions of equations, based on the well-known Newton method, are expressed by recurrences with a denominator which can vanish [17, 28, 54] as well as implicit methods for the numerical solution of differential equations [164]. Moreover, some discrete-time dynamical systems used to model the evolution of economic and financial systems, which are often

expressed by implicit recurrences $F(x_n, x_{n+1}) = 0$, assume the form of recurrences with denominator when they are expressed as $x_{n+1} = f(x_n)$ [20, 94].

1.9 Noise and Chaos: Characterization of Chaotic Behaviors

From a direct observation of a finite sequence of discrete states X_n generated by a dynamic system, it is difficult to distinguish a purely chaotic behavior (deterministic origin) from a noise effect. In the case of coexistence of these two phenomena the extraction of the chaotic signal presents major difficulties. Considering a m -dimensional dissipative system, Section 1.4.1 has described how a chaotic attractor can result from contraction of an initial volume in certain directions, and stretching toward other directions, leading to a complex folding. This process gives rise to a foliated structure with infinitely many sheets, when the (continuous or discrete) time tends toward infinity. At the limit, the figure becomes locally a Cantor set by section of the direction of contraction. An ordinary attractor A has a different behavior. Indeed it is a subset of the phase space so that, in a sufficiently small neighborhood of A , an initial volume contracts and tends asymptotically toward A . Then, in chaotic situations the corresponding strange attractor is fractal, the dimension of which is not an integer. For processes only known from time series the determination of the attractor dimension presents an interest, related to the fact that such a dimension indicates a deterministic origin for the aperiodicity observed.

In the presence of only experimental data, in the form of time series, a fundamental problem is that of discriminating the chaos from the noise, or extracting a deterministic phenomenon from the random noise. For such a purpose the power spectrum technique has limited efficiency. The notion of dimension, which can be done in several ways, is preferred. Thus one has the Kolmogorov's capacity dimension D_c , an improvement of which is the *information dimension* D_I defined from the *information entropy* [61]. The measure of correlation between points of a chaotic attractor can be made by the *correlation integral*, from which the *correlation dimension* D_{co} is defined. In general $D_{co} \leq D_I < D_c$. From an experimental signal, if D_{co} is lesser than or equal to the phase space dimension, it is likely that one has a deterministic origin. The advantage of the correlation dimension D_{co} lies in its determination which is easier to obtain than the D_c and D_I ones. The notion of *generalized information dimension* includes the aforementioned three dimensions as particular cases [74], and leads to a thermodynamic analogy. In presence of time series of only one variable, an important problem is that of the *phase space reconstitution*. A method is presented in Grassberger and

Procaccia [61], the possibilities of which are limited in presence of noise. An interesting approach for the determination of the phase space, and the distinction of a chaotic signal from a random one, is given in Sugihara and May [157] and May [96].

As mentioned earlier, a dimension not defined by an integer globally reflects an action of stretching and contraction, made in different directions on a volume of the phase space, which is reduced to an object having a smaller dimension. Such an action also appears explicitly with the notion of Liapunov's exponent. A method of calculus is given in Benettin et al., [15], and in Kaplan and Yorke [76]; it is shown that the capacity dimension is related to Liapunov exponents by a relation giving an upper bound of the dimension [61]. The Liapunov exponents can be extracted from experimental time series, after reconstitution of the phase space [41, 42, 44], the problem of noise reduction being considered in Kostelich and Yorke [89] and Hammel et al. [70]. It is worth noting that the *Kolmogorov entropy* is the sum of positive Liapunov exponents. Such exponents, as well as the use of the aforementioned dimensions, present limitations [43].

1.10 Conclusion

As discussed in Section 1.1, this chapter does not pretend to give a complete view of the scientific field presented here. It is only a guide for acquiring more extended information. Indeed, nonlinear ODEs and invertible maps have given rise to many publications. This is also the case in one-dimensional noninvertible maps, although only recently. The situation is different for the study of two-dimensional noninvertible maps, which remained a long time in an underdeveloped state. It is only in these last years that the interest in this subject has increased. One reason of this situation is the fact that more and more mathematical models of dynamical processes, belonging to different scientific fields, are related to p -dimensional noninvertible maps, $p \geq 2$. From 1964 to 1990, studies of two-dimensional noninvertible maps were made by a small number of isolated teams. Even if their number has increased since 1990, the volume of results remains very small with respect to the wide field of unknown properties. The subject of fractal basin boundaries and global properties of chaotic areas from the critical curves properties will certainly become a favorite of researchers in the near future. As for the microscopic properties, till now the results obtained concern only particular maps such as the triangular map. Continuous piecewise linear and piecewise continuous problems have given rise only to isolated results, and might be a choice of research in the future. Taking into account this situation, *a fortiori* studies of m -dimensional noninvertible maps offer an infinite

domain of investigations from the notion of r -dimensional critical sets, $r = 0, 1, \dots, m - 1$.

A class of open problem concerns the perturbation of a real two-dimensional map defined by two functions satisfying the Cauchy–Riemann conditions (cf. p. 421 of [108]). When these conditions are satisfied, the map belongs to the class of one-dimensional maps with a complex variable $z' = f(z)$, $z = x + jy$, $j^2 = -1$, studied in particular by Julia and Fatou at the beginning of the 20th century. If this perturbation leads to the nonverification of the Cauchy–Riemann conditions, then a fractal Julia set (perfect set made up of all the repulsive cycles and their limits) is destroyed. It would be interesting to identify the new fractal set generated after perturbation. Another aspect concerns the study of continuous solutions of nonlinear difference equations associated with multidimensional noninvertible maps and related problems (partial differential equations with nonlinear boundary conditions, wave propagation, etc., cf. Sharkovskij et al. [145–147]). The embedding of an m -dimensional noninvertible map into a p -dimensional invertible map, $p = m + 1, \dots, m + q$, also opens up a wide field of research [108, 124, 125]. In this case, the p -dimensional invertible map degenerates into the m -dimensional noninvertible map, when a parameter is equal to a “critical” value). Then some properties of p -dimensional map can be derived from those of the m -dimensional case. Results on invertible and noninvertible maps not defined in the whole plane (e.g., maps with denominator which can cancel, see Section 1.8) are only in an embryonic state and limited to two-dimensional maps. For p -dimensional maps, $p > 2$, focal points and prefocal curves can be extended to h -dimensional focal sets, $0 \leq h < p - 1$, and h' -dimensional prefocal sets, $1 \leq h' < p$. Such a topic induces a wide field of research from a dynamical point of view, because it also concerns map T without vanishing denominators, but such that one of the determinations T_i^{-1} of the inverse $T^{-1} = \cup_{i=1}^n T_i^{-1}$ has a vanishing denominator.

References

1. R.H. Abraham, The chaos revolution. A personal view, in *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, World Scientific Series on Nonlinear Science, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
2. R.H. Abraham and J.E. Marsden, *Foundations of Mechanics*, Addison-Wesley Publishing Co., 1985.
3. R.H. Abraham and C.D. Show, *Dynamics. The Geometry of Behavior*, Addison-Wesley Publishing Co., 1992.
4. R.H. Abraham, L. Gardini, and C. Mira, *Chaos in Discrete Dynamical Systems (A Visual Introduction in Two Dimension)*, Springer-Verlag, 1997.

5. R.H. Abraham and Y. Ueda, *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, *World Scientific Series on Nonlinear Science*, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
6. R.A. Adomaitis and I.G. Kevrekidis, Noninvertibility and the structure of basins of attraction in a model adaptive control system, *J. Non-Linear Sci.*, 1, 95–105, 1991.
7. R.A. Adomaitis, C.E. Frouzakis, and I.G. Kevrekidis, Global stability analysis of an adaptively-controlled mixing tank experiment, *Proceedings of the 1992 American Control Conference/WP4*, Chicago, 1992, pp. 1039–1043.
8. R. Allam and C. Mira, Crossroad area-dissymmetrical spring area-symmetrical spring area, and double spring area-crossroad area transition, *Int. J. Bif. Chaos*, 3 (2), 429–435, 1993.
9. A.A. Andronov, A.A. Witt, and S.E. Khaikin, *Theory of Oscillators*, Pergamon Press, 1966.
10. A.A. Andronov, E.A. Leontovich, I.I. Gordon, and A.G. Mayer, *Qualitative Theory of Dynamic Systems*, Nauka, Moscow, 1966 (in Russian).
11. A.A. Andronov, E.A. Leontovich, I.I. Gordon, and A.G. Mayer, *Bifurcation Theory of Dynamical Systems in the Plane*, Nauka, Moscow, 1967 (in Russian).
12. L.O. Barsuk, N.M. Belosludstiev, Yu.I. Neimark, and N.M. Salganskaja, Stability of a fixed point in a critical case. Bifurcations, *Radiofisica*, 11 (11), 1632–1641, 1968 (in Russian).
13. A. Barugola, J.C. Cathala, and C. Mira, Extensions of the notion of chaotic area in second order endomorphisms, *Int. J. Bif. Chaos*, 5 (3), 751–777, 1995.
14. J.R. Beddington, C.A. Free, and H. Lawton, Dynamics complexity in predator-prey models framed in difference equations, *Nature*, 255, 58–60, 1975.
15. G. Benettin, L. Galgani, A. Giorgilli, and J.M. Strelcyn, Ljapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems, *Mécanica*, 15, 9–32, 1980.
16. L. Billings and J.H. Curry, On noninvertible maps of the plane: eruptions, *Chaos*, 6, 108–119, 1996.
17. L. Billings, J.H. Curry, and E. Phipps, Symmetric functions and exact Lyapunov exponents, *Physica D*, 121, 44–64, 1998.
18. G.I. Bischi and L. Gardini, Basin fractalization due to focal points in a class of triangular maps, *Int. J. Bif. Chaos*, 7 (7), 1555–1577, 1997.
19. G.I. Bischi, L. Gardini, and C. Mira, Maps with denominator. Part 1: some generic properties, *Int. J. Bif. Chaos*, 9 (1), 119–153, 1999.
20. G.I. Bischi and A. Naimzada, Global analysis of a nonlinear model with learning, *Econ. Notes*, 26 (3), 143–174, 1997.
21. G.I. Bischi, L. Gardini, and C. Mira, New phenomena related to the presence of focal points in two-dimensional maps, *J. Anna. Math. Salesiane* (special issue), 13, 81–90, 1999.
22. G.I. Bischi, M. Kopel, and A. Naimzada, On a rent-seeking game described by a non-invertible iterated map with denominator, *Nonlinear Analysis*, T.M. and A, Proceedings of WCNA, Catania, 2000.
23. G.I. Bischi, L. Gardini, and C. Mira, Unbounded sets of attraction, *Int. J. Bif. Chaos*, 10 (9), 1437–1470, 2000.

24. G.I. Bischi, L. Gardini, and C. Mira, Maps with a vanishing denominator. A survey of some results, *Nonlinear Anal. T.M. A.* (special issue), 47, 2171–2185, 2001.
25. G.I. Bischi, M. Kopel, and A. Naimzada, On a rent-seeking game described by a non-invertible iterated map with denominator, *Nonlinear Anal. T.M. A.* (special issue Proceedings of WCNA 2000), 2001.
26. G.I. Bischi, L. Gardini, and C. Mira, Plane maps with denominator. Part 2: noninvertible maps with simple focal point, *Int. J. Bif. Chaos*, in press.
27. W.A. Brock and C.H. Hommes, A rational route to randomness, *Econometrica*, 65 (5), 1059–1095, 1997.
28. R. Brown and L.O. Chua, Clarifying Chaos: examples and counterexamples, *Int. J. Bif. Chaos*, 6 (2), 219–249, 1996.
29. J.P. Carcassès, C. Mira, C. Simo, and J.C. Tatjer, Crossroad area-spring area transition, (I) parameter plane representation, *Int. J. Bif. Chaos*, 1 (1), 183–196, 1991.
30. J.P. Carcassès, An algorithm to determine the nature and the transitions of communication areas generated by one dimensional map, *Proceeding of ECIT 91*, Lisbon, Portugal (September 15–21, 1991), World Scientific, Singapore, 1992, pp. 27–38.
31. J.P. Carcassès, Determination of different configurations of fold and flip bifurcation curves of one, or two-dimensional map, *Int. J. Bif. Chaos*, 1 (1), 869–902, 1993.
32. J.P. Carcassès, A new kind of parametric singularities and their use for the study of the bifurcation structure of a n -dimensional map, *Nonlinear Anal. Theor. Methods Appl.*, 28 (5), 917–946, 1997.
33. J.P. Carcassès and H. Kawakami, Appearance and disappearance of a dovetail structure in the parameter plane of a n -dimensional map, *Int. J. Bif. Chaos*, 9 (4), 769–783, 1999.
34. J.P. Carcassès and H. Kawakami, Existence of a cusp point on a fold bifurcation curve and stability of the associated fixed point. Case of a n -dimensional map, *Int. J. Bif. Chaos*, 9 (5), 875–894, 1999.
35. J.C. Cathala, Sur la dynamique complexe et la détermination d'une zone absorbante pour un système à données échantillonnées décrit par une récurrence du second ordre, *RAIRO Autom.*, 16 (2), 175–193, 1982.
36. J.C. Cathala, Case where the influence domain of a stable attractor of a R endomorphism is a multiply connected domain, *Proceedings of International Symposium on Iteration Theory*, Caldes de Malavella, Spain, September 20–26, 1987, World Scientific, Singapore, 1987, pp. 161–166.
37. J.C. Cathala, On some properties of absorptive areas in second order endomorphisms, *Proceedings of ECIT 89* (European Conference on Iteration Theory) Batschuns, Austria, September 1989, World Scientific, 1991, pp. 42–54.
38. J.C. Cathala, C. Mira, and H. Kawakami, Singular points with two multipliers $S_1 = +1$, $S_2 = -1$ in the bifurcation curves of maps, *Int. J. Bif. Chaos*, 2 (4), 1001–1004, 1992.
39. B. Cessac, B. Doyon, M. Quoy, and M. Samuelides, Mean-field equations, bifurcation map and route to chaos in discrete time neural networks, *Physica D*, 74, 24–44, 1994.

40. L.O. Chua, Global unfolding of Chua's circuit, *IEICE Trans. Fundam. Electron. Commun. Comp. Sci.*, 76-SA, 704–734, 1993.
41. R. Conte and M. Dubois, Lyapunov exponents of experimental systems, in *Non-Linear Evolutions*, J.P.P. Leon, Ed., World Scientific, 1988, pp. 767–780.
42. J.P. Eckmann, S. Oliffson, Kamphorst, D. Ruelle and S. Ciliberto, Lyapunov exponents from time series, *Phys. Rev.*, A34, 4971–4979, 1986.
43. J.P. Eckmann and D. Ruelle, Fundamental limitations for estimating dimensions and Ljapunov exponents in dynamical systems, *IMES*, 1989, preprint.
44. J.D. Farmer and J.J. Sidorowich, Predicting chaotic time series, *Phys. Rev. Lett.*, 59, 845–848, 1988.
45. O. Feely and L.O. Chua, Nonlinear dynamics of a class of analogue-to-digital converters, *Int. J. Bif. Chaos*, 2 (2), 325–340, 1992.
46. M.J. Feigenbaum, Quantitative universality for a class of nonlinear transformation, *J. Statist. Phys.*, 19, 25–52, 1978.
47. D. Fournier-Prunaret and N. Gicquel, Bifurcation and chaotic phenomena in an order two DPCM sytem, *Proceeding of ECCTD 95*, Istanbul, 1995.
48. C.E. Frouzakis, R.A. Adomaitis, I.G. Kevrekidis, M.P. Golden, and B.E. Ydstie, The structure of basin boundaries in a simple adaptive control system, *Proceedings of NATO 1992*, T. Bountis, Ed., Advanced Summer Institute Patras, Plenum, NY, 1992, pp. 195–210.
49. C.E. Frouzakis, R.A. Adomaitis, and I.G. Kevrekidis, An experimental and computational study of global stability for the adaptively controlled tank, *Comp. Chem. Eng.*, 20, suppl. B, 1029–1034, 1996.
50. C.E. Frouzakis, L. Gardini, I.G. Kevrekidis, G. Millerioux, and C. Mira, On some properties of invariant sets of two-dimensional noninvertible maps, *Int. J. Bif. Chaos*, 7 (6), 1167–1194, 1997.
51. L. Gardini, Global analysis and bifurcations in two-dimensional endomorphisms by use of critical lines, *Proceedings of ECIT* (European Conference on Iteration Theory) Lisbona, Portugal, September 91, World Scientific, Singapore, 1992, pp. 114–125.
52. L. Gardini, On a model of financial crisis. Critical lines as new tools of global analysis, in *Nonlinear Dynamics in Economics and Social Sciences*, F. Gori, L. Geronazzo, and M. Galeotti, Eds., Springer-Verlag, NY, 1993.
53. L. Gardini, Homoclinic bifurcations in n -dimensional endomorphisms, due to expanding periodic points, *Nonlinear Anal. Theor. Methods Appl.*, 23 (8), 1039–1089, 1994.
54. L. Gardini, G.I. Bischi, and D. Fournier-Prunaret, Basin boundaries and focal points in a map coming from Bairstow's method, *Chaos*, 9 (2), 367–380, 1999.
55. N. Gicquel and D. Fournier-Prunaret, Comportement dynamique et erreur de prédiction dans un système de transmission MICDIF, *Proceeding of GRETSI 95*, Juan les Pins, September 18–21, 1995.
56. A. Giraud, Applications des récurrences à l'étude de certains systèmes de commande, Thèse d'Ingénieur-Docteur, Faculté des Sciences de l'Université de Toulouse, 21 avril, 1969.
57. V.S. Gonchenko, D.V. Turaev, and L. Shilnikov, On models with a non-rough homoclinic Poincaré curve, *Physica D*, 62, 1–14, 1993.
58. C. Grebogi, E. Ott, and J.A. Yorke, Fractal basin boundaries, long-lived chaotic transients and unstable-stable pair bifurcations, *Physical Rev. Lett.*, 50 (13), 935–938, 1983.

59. C. Grebogi, E. Ott, and J.A. Yorke, Chaos, strange attractors, and fractal basin boudaries in nonlinear dynamics, *Science*, 238, 632–638, 1987.
60. C. Grebogi, S.M. Hammel, J.A. Yorke, and T. Sauer, Shadowing of physical trajectories in chaotic dynamics: containment and refinement, *Phys. Rev. Lett.*, 65 (13), 1527–1530, 1990.
61. P. Grassberger and I. Procaccia, Measuring the strangeness of strange attractors, *Physica 9D*, 189–208, 1983.
62. J. Guckenheimer, G.F. Oster, and A. Ipaktchi, The dynamics of density depended population models, *J. Math. Biol.*, 4, 101–147, 1977.
63. J. Guckenheimer, The bifurcations of quadratic functions, *N.Y. Acad. Sci.*, 75/1, 343–347, 1980.
64. J. Guckenheimer and P. Holmes, *Nonlinear Oscillator, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.
65. J. Guckenheimer and R.F. Williams, Structural stability of the attractor in a Lorenz model, *Inst. Hautes Etudes Sci. Publ. Math.*, 50, 59–72, 1979.
66. I. Gumowski and C. Mira, Solutions chaotiques bornées d’une récurrence ou transformation ponctuelle du second ordre à inverse non unique, *C. R. Acad. Sci. Paris, Sér. A*, 285, 477–480, 1977 (in French).
67. I. Gumowski and C. Mira, Bifurcation déstabilisant une solution chaotique d’un endomorphisme du 2nd ordre, *C. R. Acad. Sci. Paris, Sér. A*, 286, 427–431, 1978 (in French).
68. I. Gumowski and C. Mira, *Dynamique chaotique. Transformations ponctuelles. Transition ordre désordre*, Cépadues Editions Toulouse, 1980.
69. I. Gumowski and C. Mira, Recurrences and Discrete Dynamic Systems, Lecture Notes in Mathematics, 809, Springer Verlag, 1980.
70. S.M. Hammel, J.A. Yorke, and C. Grebogi, Do numerical orbits of chaotic dyanmical processes represent true orbits? *J. Complexity*, 3, 136–140, 1987.
71. M. Hasler, Synchronization principles and applications, in *Circuits and Systems Tutorial*, C. Toumazou, Ed., IEEE ISCAS’94, London, 1994, pp. 314–327.
72. C. Hayashi, *Nonlinear Oscillations in Physical Systems*, McGraw-Hill, New York, 1964.
73. Ch. Hayashi, Y. Ueda, and H. Kawakami, Periodic solutions of Duffing’s equation with reference to doubly asymptotic solutions, *Proceedings of 5th International Conference on Non-Linear Oscillations*, Kiev, 1969, pp. 235–255.
74. H.G.E. Hentschel and I. Procaccia, The infinite number of generalized dimensions of fractals and strange attractors, *Physica 8D*, 435–444, 1983.
75. P. Holmes and R.F. Wiliams, Knotted periodic orbits in suspensions of smale’s horseshoe: Torus knots and bifurcation sequences, *Arch. Rat. Mech. Anal.*, 90 (2), 115–195, 1985.
76. J.L. Kaplan and J.A. Yorke, Chaotic behavior of multidimensional difference equations, in *Functional Difference Equations, Lecture Notes in Math*, 730, pp. 204–228, 1979.
77. A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, 1995.
78. H. Kawakami, *C.R. Acad. Sci. Paris*, 293 (Sér. I), 401–403, 1981.
79. H. Kawakami, Bifurcation of periodic responses in forced dynamics nonlinear circuits: computation of bifurcation values of the system parameters, *IEEE Trans. Circuits Syst.*, CAS-31 (3), 248–260, 1984.

80. H. Kawakami and K. Kobayashi, Computer experiments on chaotic solutions of $x(t+2) - ax(t+1) - x^2(t) = b$, *Bull. Fac. Eng., Tokushima Univ.*, 16, 29–46, 1979.
81. H. Kawakami and J. Matsuo, Bifurcation of doubly asymptotic motions in nonlinear systems, *IEICE Trans.*, J65-A (7), 647–654, 1982.
82. H. Kawakami and M. Sekai, Synchronization Phenomena Observed in a Duffing–Rayleigh Equation, Publication NLP83-37, Tokushima University, 1983 (in Japanese).
83. H. Kawakami and M. Sekai, Synchronization Phenomena and Hopf Bifurcation, Publication NLP84-7, Tokushima University, 1984, pp. 25–33 (in Japanese).
84. H. Kawakami and M. Sekai, Synchronization and chaos in a Duffing-Rayleigh oscillator, in *Asymptotic Methods of Mathematical Physics*, (collection of scientific papers), Kiev, Naukova Dumka, 1988, pp. 83–92.
85. Y. Kevrekidis, R.A. Adomaitis, and R. de la Llave, Predicting the Complexity of Disconnected Basins of Attraction for a Noninvertible System, Systems Research Center, Technical Report #TR91-41, University of Maryland, 1991.
86. Y. Kevrekidis, R.A. Adomaitis, and R. de la Llave, On the Dynamics and Global Stability Characteristics of Adaptive Control Systems, Systems Research Center, Technical Report #TR91-100, University of Maryland, 1991.
87. H. Khammary, Sur le comportement complexe des oscillations forcées de systèmes non linéaires soumis à une excitation sinusoïdale (in French), Thèse de Doctorat en génie Electrique de l'Ecole Nationale d'Ingénieurs de Tunis (12 Février 1999).
88. H. Khammary, J.P. Carcasses, and C. Mira, Higher harmonic resonances in a RLC circuit, to be published in *Int. J. Bif. Chaos*, 2006.
89. E.J. Kostelich and J.A. Yorke, Noise reduction in dynamical systems, *Phys. Rev.*, A38, 1649–1652, 1988.
90. T.Y. Li and J.A. Yorke, Exploring chaos on an interval, in *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, *World Scientific Series on Nonlinear Science*, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
91. V. Maistrenko, Yu. Maistrenko, and I. Sushko, Noninvertible two-dimensional maps arising in radiophysics, *Int. J. Bif. Chaos*, 4 (2), 383–400, 1994.
92. V. Maistrenko, Yu. Maistrenko, and I. Sushko, Period adding phenomenon in piecewise linear endomorphisms arising from electronic systems, *Proceedings of Workshop NDES 1994*, Krakow, Poland, 1994, pp. 39–44.
93. Yu. Maistrenko and I. Sushko, Bifurcation phenomena in auto oscillations systems with delay leading to two-dimensional piecewise linear maps, *Proceedings of Workshop NDES 1995*, Dublin, Ireland, 1995, pp. 83–87.
94. R. Marimon and S. Sunder, Expectations and learning under alternative monetary regimes: an experimental approach, *Econ. Theor.*, 4, 131–162, 1994.
95. J.R. Marotto, Snap-back repellers imply chaos in R^n , *J. Math. Anal. Appl.*, 63, 199–223, 1978.
96. R.M. May, Le chaos en Biologie, *La Recherche*, 22 (232), 589–598, 1991 (in French).
97. N. Metropolis, M.L. Stein, and P.R. Stein, On limit finite sets for the transformation of the unit interval, *J. Comb. Theor.*, (A), 15, 25–44, 1973.

98. G. Millérioux and C. Mira, Homoclinic and heteroclinic situations specific to two-dimensional noninvertible maps, *Int. J. Bif. Chaos*, 7 (1), 39–70, 1997.
99. G. Millérioux and C. Mira, Coding scheme based on chaos synchronization from noninvertible maps, *Int. J. Bif. Chaos*, 8 (8), 1812–1824, 1998.
100. J. Milnor and R. Thurston, *On Iterated Maps of the Interval*, Princeton University Press, 1977, unpublished notes.
101. C. Mira, Etude d'un premier cas d'exception pour une récurrence, ou transformation ponctuelle, du deuxième ordre, *C. R. Acad. Sci. Paris, Sér. A*, 269, 1006–1009, 1969.
102. C. Mira, Etude d'un second cas d'exception pour une récurrence, ou transformation ponctuelle, du deuxième ordre, *C. R. Acad. Sci. Paris, Sér. A*, 270, 332–335, 1970.
103. C. Mira, Sur les cas d'exception d'une récurrence, ou transformation ponctuelle, du deuxième ordre, *C. R. Acad. Sci. Paris, Sér. A*, 270, 466–469, 1970.
104. C. Mira, Accumulations de bifurcations et structures boîtes-emboîtées dans les récurrences et transformations ponctuelles, in *Proceedings of the 7th International Conference on Nonlinear Oscillations*, Berlin, September 1975, Akademik Verlag, Berlin 1977, Band I2, 1975, pp. 81–93.
105. C. Mira, Sur la notion de frontière floue de stabilité, in *Proceedings of the 3rd Brazilian Congress of Mechanical Engineering*, Rio de Janeiro, December 1975, D4, 1975, pp. 905–918.
106. C. Mira, Frontière floue séparant les domaines d'attraction de deux attracteurs, *C. R. Acad. Sci. Paris, Sér. A*, 288, 591–594, 1979.
107. C. Mira, Complex dynamics in two-dimensional endomorphisms, *Nonlinear Anal., T.M. A.*, 4 (6), 1167–1187, 1980.
108. C. Mira, *Chaotic Dynamics. From the One-Dimensional Endomorphism to the Two-Dimensional Diffeomorphism*, World Scientific, Singapore, 1987, 450 pp.
109. C. Mira, *Systèmes asservis non linéaires*, Hermès, Paris, 1990, 425 pp. (in French).
110. C. Mira, On some bifurcations structures occurring in nonlinear dynamics, in *Proceedings of the Second Symposium on Nonlinear Theory and Its Applications (NOLTA 91)*, Fukuoka, July 1991, pp. 107–114.
111. C. Mira, Some historical aspects of nonlinear dynamics. Possible trends for the future, Double publication: (1) *Int. J. Bif. Chaos*, 7 (9 and 10), 2145–2174, 1997. (2) *J. Franklin Inst.*, 334B (5/6), 1075–1113, 1997.
112. C. Mira, Chaos and fractal properties induced by noninvertibility of models in the form of maps, *Chaos Solitons Fractals*, (11), 251–262, 2000.
113. C. Mira and J.C. Roubellat, Cas où le domaine de stabilité d'un ensemble limite attractif d'une récurrence n'est pas simplement connexe, *C. R. Acad. Sci. Paris, Sér. A*, 268, 1657–1660, 1969.
114. C. Mira and J.P. Carcassès, On the crossroad area–saddle area and crossroad area–spring area transitions, *Int. J. Bif. Chaos*, 1 (3), 641–655, 1991.
115. C. Mira, J.P. Carcassès, C. Simo, and J.C. Tatjer, Crossroad area–spring area transition. (II) Foliated parametric representation, *Int. J. Bif. Chaos*, 1 (2), 339–348, 1991.
116. C. Mira and H. Kawakami, Qualitative modifications of the lip bifurcation structure, in *Proceedings of the European Conference on Iteration Theory, ECIT'92*,

- Batschuns, September 13–19, 1992 (Austria), Förg-Rob et al., Eds., World Scientific, 1992, pp. 199–203.
117. C. Mira and I. Djellit, Bifurcations structure in a model of frequency modulated CO₂ laser, *Int. J. Bif. Chaos*, 3 (1), 97–129, 1993.
 118. C. Mira and M. Qriouet, On a ‘Crossroad area–spring area’ transition occurring in a Duffing–Rayleigh equation with periodical excitation, *Int. J. Bif. Chaos*, 3 (4), 1029–1037, 1993.
 119. C. Mira, H. Kawakami, and R. Allam, The dovetail bifurcation structure and its qualitative changes, *Int. J. Bif. Chaos*, 3 (4), 1029–1037, 1993.
 120. C. Mira, D. Fournier-Prunaret, L. Gardini, H. Kawakami, and J.C. Cathala, Basin bifurcations of two-dimensional noninvertible maps: fractalization of basins, *Int. J. Bif. Chaos*, 4 (2), 343–381, 1994.
 121. C. Mira, L. Gardini, A. Barugola, and J.C. Cathala, *Chaotic Dynamics in Two-Dimensional Noninvertible Maps*, World Scientific Series on Nonlinear Sciences, Ser. A, Vol. 20, 1996, 630 pp.
 122. C. Mira, G. Millerioux, J.P. Carcasses, and L. Gardini, Plane foliation of two-dimensional noninvertible maps, *Int. J. Bif. Chaos*, 6 (8), 1439–1462, 1996.
 123. C. Mira, H. Kawakami, and M. Touzani-Qriouet, Bifurcations structures generated by the non-autonomous Duffing equation, *Int. J. Bif. Chaos*, 9 (7), 1363–1379, 1999.
 124. C. Mira, H. Abdel-Basset, and H. El-Hamouly, Implicit approximation of a stable saddle manifold generated by a two-dimensional quadratic map, *Int. J. Bif. Chaos*, 9 (8), 1535–1547, 1999.
 125. C. Mira and C. Gracio, On the embedding of a $(p - 1)$ -dimensional non invertible map into a p -dimensional invertible map ($p = 2, 3$), *Int. J. Bif. Chaos*, in press.
 126. P.J. Myrberg, Iteration von Quadratwurzeloperationen. I, *Ann. Acad. Sci. Fenn., Ser. A*, 256, 1–10, 1958.
 127. P.J. Myrberg, Iteration von Quadratwurzeloperationen. II, *Ann. Acad. Sci. Fenn., Ser. A*, 268, 1–10, 1959.
 128. P.J. Myrberg, Sur l’itération des polynômes réels quadratiques, *J. Math. Pures Appl.*, 41 (9), 339–351, 1962.
 129. P.J. Myrberg, Iteration von Quadratwurzeloperationen. III, *Ann. Acad. Sci. Fenn., Ser. A*, 336, 1–10, 1963.
 130. Yu.I. Neimark, *The Method of Point Mappings in the Theory of Non Linear Oscillations*, Nauka, Moscow, 1972 (in Russian).
 131. S.E. Newhouse, Diffeomorphisms with infinitely many sinks, *Topology*, 13, 9–18, 1974.
 132. S.E. Newhouse, The abundance of wild hyperbolic sets and non-smooth stable sets for diffeomorphisms, *Publ. Math. IHES*, 50, 101–151, 1979.
 133. M. Ogorzalek and H. Dedieu, Chaos control technics for signal processing, in *Proceedings of 1995 IEEE Workshop on Nonlinear Signal and Image Processing*, Neos Marmaras, Halkidiki, Greece, 1995.
 134. E. Ott, C. Grebogi, and J.A. Yorke, Controlling chaos, *Physi. Rev. Lett.*, 64 (11), 1196–1199, 1990.
 135. T.S. Parker and L.O. Chua, *Practical Numerical Algorithms for Chaotic Systems*, Springer-Verlag, 1989.

136. L.M. Pecora and T.L. Carol, Synchronization in chaotic systems, *Phys. Rev. Lett.*, 64, 821–824, 1990.
137. M. Quoy, B. Cessac, B. Doyon, and M. Samuelides, Dynamical behaviour of neural networks with discrete time dynamics, *Neural Network World*, 3 (6), 845–848, 1993.
138. M. Qriouet and C. Mira, Fractional harmonic synchronization in the Duffing-Rayleigh differential equation, *Int. J. Bif. Chaos*, 4 (2), 411–426, 1994.
139. Rico-Martinez, R. Adomaitis, and Y. Kevrekidis, Noninvertibility in neural networks, *Proceedings of the 1993 IEEE International Conference on Neural Networks*, San Francisco, 1993, pp. 382–386.
140. C. Robinson, *Dynamical Systems*, CRC Press, 1995.
141. O.E. Rössler, Chaos, hyperchaos and the double perspective, in *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, World Scientific Series on Nonlinear Science, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
142. A.N. Sharkovskij, Coexistence of cycles of a continuous map of a line into itself, *Ukrain. Mat. J.*, 16 (1), 61–71, 1964.
143. A.N. Sharkovskij, Problem of isomorphism of dynamical systems, in *Proceeding of the 5th International Conference on Nonlinear Oscillations*, Vol. 2, Kiev, 1969, pp. 541–544.
144. A.N. Sharkovskij, On some properties of discrete dynamical systems, in *Proceedings of Théorie de l'iteration et ses applications*, Toulouse, Ed., CNRS, 1982, pp. 153–158.
145. A.N. Sharkovskij and E. Yu. Romanenko, Ideal turbulence: attractors of deterministic systems may lie in the space of random fields, *Int. J. Bif. Chaos*, 2 (1), 31–36, 1992.
146. A.N. Sharkovskij, Yu.L. Maistrenko, and E.Yu. Romanenko, *Difference Equations and Their Applications*, Series Mathematics and Its Applications, Kluwer Academic Publishers, 1993, 358 pp.
147. A.N. Sharkovsky, Yu.L. Maistrenko, Ph. Deregel, and L.O. Chua, Dry turbulence from a time-delayed Chua's circuit, *J. Circuits Syst. Comp.*, 3 (2), 645–668, 1993.
148. L.P. Shilnikov, Strange attractors and dynamical models, *J. Circuits Syst. Comp.*, 3, 1–10, 1993.
149. L.P. Shilnikov, Mathematical problem of nonlinear dynamics: a tutorial, *Int. J. Bif. Chaos*, 7 (9), 1953–2001, 1997.
150. L. Shilnikov, A. Shilnikov, D. Turaev, and L. Chua, *Methods of Qualitative Theory in Nonlinear Dynamic*, Part I (see also Part II, 2001). World Scientific, Singapore, 1998.
151. T. Shinbrot, C. Grebogi, E. Ott, and J.A. Yorke, Using small perturbations to control chaos, *Nature*, 363, 411–417, 1993.
152. S. Smale, Morse inequalities for a dynamical system, *Bull. Am. Math. Soc.*, 66, 43–49, 1960.
153. S. Smale, Diffeomorphisms with many periodic points, in *Differential Combinatorial Topology*, S.S. Cairns, Ed., Princeton University Press, 1963, pp. 63–80.
154. S. Smale, Structurally stable systems are not dense, *Am. J. Math.*, 88, 491–496, 1966.

155. S. Smale, Differentiable dynamical systems, *Bull. Am. Math. Soc.*, 73, 747–817, 1967.
156. S. Smale, Finding a horseshoe on the beaches of Rio, in *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, World Scientific Series on Nonlinear Science, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
157. G. Sugihara and R.M. May, *Nature*, (344), 734–740, 1990.
158. M. Touzani-Qriouet and C. Mira, Reducible fractional harmonics generated by the nonautonomous Duffing–Rayleigh equation. Pockets of reducible hartmonics and Arnold’s tongues, *Int. J. Bif. Chaos*, 10 (6), 1345–1366, 2000.
159. Y. Ueda, *The Road to Chaos*, Aerial Press, Inc., Santa Cruz, USA, 1992.
160. Y. Ueda, Strange attractors and the origin of chaos, in *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, World Scientific Series on Nonlinear Science, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
161. Y. Ueda, My encounter with chaos, in *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, World Scientific Series on Nonlinear Science, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
162. Y. Ueda, Reflections on the origin of the broken-egg chaotic attractor, in *The Chaos Avant-Garde. Memories of the Early Days of Chaos Theory*, World Scientific Series on Nonlinear Science, L.O. Chua, Ed., Series A, Vol. 39, 2000, 219 pp.
163. S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, 1990.
164. H.C. Yee and P.K. Sweby, Global asymptotic behavior of iterative implicit schemes, *Int. J. Bif. Chaos*, 4 (6), 1579–1611, 1994.
165. T. Yoshinaga, H. Kitajima, H. Kawakami, and C. Mira, A method to calculate homoclinic points of a two dimensional noninvertible map, *IEICE Trans. Fundam.*, E80-A (9), 1560–1566, 1997.

2

Tools for Ordinary Differential Equations Analysis

W. Perruquetti

CONTENTS

2.1	Introduction	46
2.1.1	Biology	47
2.1.2	Chemistry	47
2.1.3	Electricity	48
2.1.4	Electrical Motors	49
2.1.5	Mechanics	49
2.2	First-Order Differential Equation	50
2.2.1	Notion of Solution	52
2.2.1.1	Phase Portrait Solution	52
2.2.1.2	Existence	53
2.2.1.3	Extension, Uniqueness, and Global Solution	56
2.2.1.4	Dependence of the Initial Conditions	58
2.2.2	Classification	58
2.2.2.1	Autonomous Linear Case	59
2.2.2.2	Nonlinear Autonomous Case	60
2.2.3	Flow	61
2.3	Some Characterizations of Behaviors	62
2.3.1	Remarkable Sets	62
2.3.1.1	Equilibrium Point	62
2.3.1.2	Orbits: Periodic, Closed, Homoclinic, and Heteroclinic	65
2.3.2	Properties	67
2.3.2.1	Invariance	67
2.3.2.2	Liapunov Stability	68
2.3.2.3	Attractivity	70
2.3.2.4	Asymptotic Stability	72
2.3.2.5	Exponential Stability	74
2.4	Autonomous Linear Case	75

2.4.1	Formal Computation of Solution	75
2.4.2	Stability Conditions	77
2.5	Behavior Studies: Local Results	79
2.5.1	Structural Stability	79
2.5.1.1	Whitney Distance	79
2.5.1.2	Equivalence	79
2.5.1.3	Structural Stability	80
2.5.2	From Linear Model to Nonlinear Model	80
2.5.2.1	Structural Stability Theorem and Consequences	81
2.5.2.2	Local Structure of Solution Within a Neighborhood of an Equilibrium Point	82
2.5.2.3	Local Structure of the Solutions in the Neighborhood of a Closed Orbit	86
2.6	Bifurcation and Chaos	88
2.6.1	Parameter Dependence	88
2.6.2	Local Bifurcation Locale with Codimension 1	91
2.6.2.1	Subcritical or Saddle–Node	92
2.6.2.2	Transcritical Bifurcation	93
2.6.2.3	Supercritical	94
2.6.3	Chaos	95
	References	98

2.1 Introduction

Many physical modeling activities of the 16th century were conducted within the framework of infinitesimal calculus (nowadays known as differential calculus). Indeed, these models are relations between variables which are functions of a special variable named “time” and their derivatives with respect to this time variable: these relations are *ordinary differential equations* (ODEs). Isaac Newton (1642–1727) in his 1687 memoir titled *Philosophiae naturalis principiae mathematica* wrote: “*Data aequatione quotcunque fluentes quantitates involvente fluxiones invenire et vice versa*” (he is underlying the fundamental role played by ODE). Since then, many physical process were described using ODEs (e.g., in the 17th century Euler–Lagrange equations were used to describe mechanical systems).

ODE models are used in several fields ranging from biology to mechanics.

2.1.1 Biology

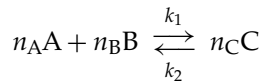
Consider bacterium are growing on a substrate in a Petri box. Let x be the bacterium number, a simplified model, called the *logistic model*, is:

$$\frac{dx}{dt} = ax(x_{\max} - x) \quad (2.1)$$

where a is a strictly positive constant and $x_{\max}$ is the maximum number of bacteria which can live in the box. Indeed, when there are few bacteria: $\dot{x} \sim ax$ (exponential growth) and when x is close to $x_{\max}$, the growth is reduced since $\dot{x} \sim 0$. Another example is the Volterra model for predator-prey co-evolution (see Example 3 in Section 2.3.1).

2.1.2 Chemistry

Different balance sheets (of matter, thermodynamics) can, when reduced to their lowest terms, be expressed using ODEs. For example, let us consider a tank filled with two chemicals A and B, whose concentrations are, respectively, c_A and c_B , with respective flows of u_1 and u_2 calculated by the use of two pumps. In this vat, a mixer homogenizes both products, which react according to:



where n_A , n_B , and n_C are, respectively, the stoichiometric coefficients of A, B, and C. The mixture is flowing off the vat through an aperture of section s to the base of this vat (whose section is $S = 1 \text{ m}^2$). The balance sheet of matter conducts, using the Bernoulli relation, is:

$$S \frac{dh}{dt} = u_1 + u_2 - \sqrt{2sgh}$$

where h is the height of mixture in the tank and g is the gravitational constant (9.81 m sec^{-2}). Laws of the kinetics give the relation (under the hypothesis of a second-rate kinetics):

$$v_{\text{cin}} = -k_1 c_A c_B + k_2 c_C^2$$

Therefore:

$$\begin{aligned}\frac{d(hc_A)}{dt} &= u_1 c_{A0} - \sqrt{2sg} hc_A - n_A v_{\text{cin}} h \\ \frac{d(hc_B)}{dt} &= u_2 c_{B0} - \sqrt{2sg} hc_B - n_B v_{\text{cin}} h \\ \frac{d(hc_C)}{dt} &= -\sqrt{2sg} hc_C + n_C v_{\text{cin}} h\end{aligned}$$

with $c_{A0} = c_A(\text{entering})$ and $c_{B0} = c_B(\text{entering})$. Denoting the state vector by $x = (h, hc_A, hc_B, hc_C)^T$, the model reads as:

$$\begin{cases} \dot{x}_1 = u_1 + u_2 - \sqrt{2sg} \sqrt{x_1} \\ \dot{x}_2 = u_1 c_{A0} - \sqrt{\frac{2sg}{x_1}} x_2 - n_A \frac{(-k_1 x_2 x_3 + k_2 x_4^2)}{x_1} \\ \dot{x}_3 = u_2 c_{B0} - \sqrt{\frac{2sg}{x_1}} x_3 - n_B \frac{(-k_1 x_2 x_3 + k_2 x_4^2)}{x_1} \\ \dot{x}_4 = -\sqrt{\frac{2sg}{x_1}} x_4 + n_C \frac{(-k_1 x_2 x_3 + k_2 x_4^2)}{x_1} \end{cases} \quad (2.2)$$

2.1.3 Electricity

An electrical system is made of a resistor R , an inductance L , and a capacitor C , each in a branch of a triangle. Let us note, respectively, i_X and v_X to be the current and the voltage in the branch where X is. Assuming that L and C are linear and that only the resistor R is nonlinear but satisfies the generalized Ohm's law ($v_R = f(i_R)$); the Kirchhoff laws leads to:

$$\begin{cases} L \frac{di_L}{dt} = v_L = v_C - f(i_L) \\ C \frac{dv_C}{dt} = i_C = -i_L \end{cases} \quad (2.3)$$

If in this ODE, called *Liénard equation*, one considers the particular case where $f(x) = (x^3 - 2\mu x)$, then one gets the *Van der Pol equation*.

Another famous example is the Chua's circuit: a nonlinear resistor R_{nl} , satisfying the generalized Ohm's law:

$$i = f(v) = G_b + \frac{1}{2}(G_a - G_b)\{|v_1 + E| - |v_1 - E|\}$$

in parallel with a capacitor C_1 coupled through a resistor R to an inductance L with linear resistor R_0 in parallel with an other capacitor C_2 . Denoting

the current of the inductance by i and the voltages of the capacitors C_1 and C_2 by, respectively, v_1 and v_2 arrive at the following model:

$$\begin{cases} C_1 \frac{dv_1}{dt} = \frac{v_2 - v_1}{R} - f(v_1) \\ C_2 \frac{dv_1}{dt} = \frac{v_1 - v_2}{R} + i \\ L \frac{di}{dt} = v_2 + R_0 i \end{cases} \quad (2.4)$$

2.1.4 Electrical Motors

For a stepper motor with n pairs of teeth, the electromagnetic balance (in the dq frame called Park frame) is:

$$\begin{aligned} L_d \frac{di_d}{dt} &= v_d - Ri_d + nL_q \omega i_q \\ L_q \frac{di_q}{dt} &= v_q - Ri_q - nL_d \omega i_d - nmi_r \omega \\ C_{em} &= n(L_d - L_q)i_d i_q + nmi_r i_q + K_d \sin(n\theta) \end{aligned}$$

where m and i_r are, respectively, the inductance and the fictitious rotor current, leading to the constant flux mi_r (permanent magnet); i_d, i_q, v_d, v_q are the currents and voltages in the dq frame, respectively. The mechanical balance is:

$$\begin{aligned} \frac{d\theta}{dt} &= \omega \\ J \frac{d\omega}{dt} &= C_{em} - C_{load} \end{aligned}$$

2.1.5 Mechanics

If a mechanical system is made of n links connected by means of perfect joints (without friction), one will have the position of the system which will depend on n independent parameters (generalized coordinates denoted by $q_1, \dots, q_n$). To write the *Euler-Lagrange equations*, the *lagrangian* must be determined (difference between the kinetic energy and the potential energy):

$$\mathcal{L} = \mathcal{E}_c - \mathcal{E}_p \quad (2.5)$$

the elementary work of each internal and external forces D_i , as well as the work of friction forces:

$$-\frac{\partial D}{\partial \dot{q}_i} dq_i,$$

gives some dissipative energy D . One thus obtains the well-known Euler–Lagrange system:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} + \frac{\partial D}{\partial \dot{q}_i} = D_i \quad (2.6)$$

Let us note that the kinetic energy $\mathcal{E}_c = (1/2)\dot{q}^T M(q)\dot{q}$, with $M(q)$ an $n \times n$ positive definite matrix, depends on q_i and their derivatives $\dot{q}_i$, whereas the potential energy $\mathcal{E}_p$ depends only on q_i . For a pendulum (θ being the angle between the rope and the vertical position) one gets $\mathcal{L} = (1/2)ml^2\dot{\theta}^2 - mgl(1 - \cos(\theta))$. Neglecting the friction terms one gets:

$$\ddot{\theta} = -\frac{g}{l} \sin(\theta) \quad (2.7)$$

When dealing with such models, several questions arise: what do we mean by a solution to such an ODE? Do there exist conditions ensuring the existence of such a solution? Some results will be discussed in Section 2.2. Beyond these existence conditions, one can ask about the qualitative properties of such solutions: can we characterize asymptotic behavior (see Section 2.3 devoted equilibrium points, limit cycle and strange attractor)? Section 2.4 deals with the particular case of linear systems which, along with Section 2.5, presents tools to analyze and characterize the asymptotic behavior of solution near such sets. Lastly, almost every physical system involves in its model some parameters which, when varying, may modify the qualitative properties of the solutions: which is the scope of Section 2.6.

2.2 First-Order Differential Equation

An *implicit ODE* is of the following form:

$$F \left(t, y, \frac{dy}{dt}, \dots, \frac{d^k y}{dt^k} \right) = 0, \quad y \in \mathbb{R}^m \quad (2.8)$$

with F defined on an open set of $\mathbb{R} \times \mathbb{R}^{m(k+1)}$ and taking a value in $\mathbb{R}^m$. The *order* of the ODE is the integer k which is the higher order derivative in the

relation (2.8). Let us note that (2.1), (2.2), and (2.7) are of order, respectively, 1, 1, and 2. The implicit function theorem ensures that the m relations in (2.8) can be expressed (at least locally) as an *explicit ODE* as:

$$\frac{d^k y}{dt^k} = G \left(t, y, \frac{dy}{dt}, \dots, \frac{d^{k-1}y}{dt^{k-1}} \right) \quad (2.9)$$

as soon as:

$$\det(J_F) \neq 0$$

where J_F is the jacobian matrix of function F ; this is a matrix with entries

$$a_{ij} = \frac{\partial F_i}{\partial \left(d^k x_j / dt^k \right)} \quad (i, j) \in \{1, \dots, m\}^2$$

Note that when the variable y in the implicit ODE (2.8) belongs to a more general set than the cartesian product of open set in $\mathbb{R}$, then letting

$$x = \left(y, \frac{dy}{dt}, \dots, \frac{d^{k-1}y}{dt^{k-1}} \right)^T$$

the explicit ODE (2.9) can be written in the form:

$$\frac{dx}{dt} = f(t, x), \quad t \in \mathcal{I}, \quad x \in \mathcal{X} \quad (2.10)$$

In this expression: $t \in \mathcal{I} \subset \mathbb{R}$ represents the *time variable* and $\mathcal{X}$ is the *state space*.¹ For practical reasons, the state space may be bounded in order to take into account physical limitations. In general, the state space is a differentiable manifold. When the vector x contains a variable and its successive derivatives, $\mathcal{X}$ is then called *phase space*. However, some authors (p. 11 of [2]) use the two designations without discrimination. The vector $x \in \mathcal{X}$ is the *state vector* of (2.10) (sometimes the phase vector according to the situation). In practice, it contains a sufficient number of variables useful to describe the time evolution of the process; $x(t)$ is the instantaneous state at time t and $f : \mathcal{I} \times \mathcal{X} \rightarrow T\mathcal{X}$ (tangent space), $(t, x) \mapsto f(t, x)$, is the *vector field*. To simplify the rest of the presentation, we will consider the particular case where $\mathcal{I} \times \mathcal{X}$ is an open of $\mathbb{R}^{n+1}$ and $T\mathcal{X}$ is $\mathbb{R}^n$.

¹Words used in the field of automatic control.

2.2.1 Notion of Solution

When speaking about solution, one has to state precisely the associated problem: for ODE, there exist a boundary problem² and an *initial condition problem* (called *the Cauchy Problem*, CP):

$$(CP): \begin{cases} \text{"Do there exist a function:} \\ \phi : \mathcal{I} \subset \mathbb{R} \rightarrow \mathcal{X} \subset \mathbb{R}^n \\ t \mapsto \phi(t) \\ \text{satisfying (2.10) and the given initial condition: } \phi(t_0) = x_0? \end{cases}$$

We are looking for a sufficiently smooth function of time $\phi : t \mapsto \phi(t)$, whose time derivative is (for almost all times³) the same as the value of the vector field evaluated at the same instant and at the location given by this function $x = \phi(t)$. If $f(u, \phi(u))$ is measurable,⁴ then one can rewrite $\phi(t)$ in the following form:

$$\phi(t) = \phi(t_0) + \int_{t_0}^t f(v, \phi(v)) \, dv \quad (2.11)$$

the integral has to be understood in the Lebesgue sense and this, even if $t \mapsto f(t, \cdot)$ is not continuous with respect to t [which may be useful when dealing with $\dot{x} = g(t, x, u)$ because a discontinuous feedback of the form $u = u(t)$ can be used]. Thus, we will look for functions which are at least absolutely continuous⁵ with respect to time.

2.2.1.1 Phase Portrait Solution

DEFINITION 1 A solution of (2.10) originating from x_0 at t_0 is any absolutely continuous function ϕ defined on a non-empty set $\mathcal{I}(t_0, x_0) \subset \mathcal{I} \subset \mathbb{R}$ which contains t_0 :

$$\begin{aligned} \phi : \mathcal{I}(t_0, x_0) \subset \mathcal{I} \subset \mathbb{R} &\rightarrow \mathcal{X} \subset \mathbb{R}^n \\ t &\mapsto \phi(t; t_0, x_0), \end{aligned}$$

²Similar to the CP, for which the data of initial condition is replaced by n data $\phi_{\sigma(i)}(t_i)$ at given times $t_i, i \in \mathbb{N} = \{1, \dots, n\}, \sigma : \mathbb{N} \rightarrow \mathbb{N}$.

³This is to say for all times $t \in \mathcal{T} \setminus \mathcal{M}$, with $\mathcal{M}$ a set of zero measure, using the following notation $\mathcal{T} \setminus \mathcal{M} = \{x \in \mathcal{T} : x \notin \mathcal{M}\}$.

⁴This holds if, for x fixed, $t \mapsto f(t, x)$ is measurable and for t fixed, $x \mapsto f(t, x)$ is continuous.

⁵ $\phi : [\alpha, \beta] \mapsto \mathbb{R}^n$ is *absolutely continuous* if $\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0 : \forall \{\alpha_i, \beta_i\}_{i \in \{1, \dots, n\}}, [\alpha_i, \beta_i] \subset [\alpha, \beta], \sum_{i=1}^n (\beta_i - \alpha_i) \leq \delta(\varepsilon) \Rightarrow \sum_{i=1}^n \|\phi(\beta_i) - \phi(\alpha_i)\| \leq \varepsilon$. ϕ is absolutely continuous if and only if there exists a Lebesgue integrable function which is the derivative of ϕ almost everywhere.

in short denoted by $\phi(t)$, which satisfies (2.11) for all $t \in \mathcal{I}(t_0, x_0)$ (or equivalently: $\dot{\phi} = f(t, \phi(t))$ for almost all time in $\mathcal{I}(t_0, x_0)$) and such that $\phi(t_0) = x_0$.

Example 1

Using separation of variables, the logistic equation (2.1) becomes:

$$\frac{dx}{ax(x_{\max} - x)} = dt$$

which can be used to obtain a solution to the CP (2.1), $x(0) = x_0$:

$$\begin{aligned} \phi : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \phi(t; 0, x_0) = \frac{x_0 x_{\max}}{x_0 + e^{-ax_{\max}t}(x_{\max} - x_0)} \end{aligned} \quad (2.12)$$

DEFINITION 2 A solution of (2.10) can be viewed:

- Either in the extended state space $\mathcal{I} \times \mathcal{X}$ named the space of motion, in that case one is talking about motion or trajectory
- Or in the state space $\mathcal{X}$, in which case one is talking about *orbit*.

The set of all possible orbit oriented with respect to time is called the *phase portrait*.

Usually, when drawing the phase portrait, only accumulating sets are drawn as time tends to $\pm\infty$. For example, for the following system:

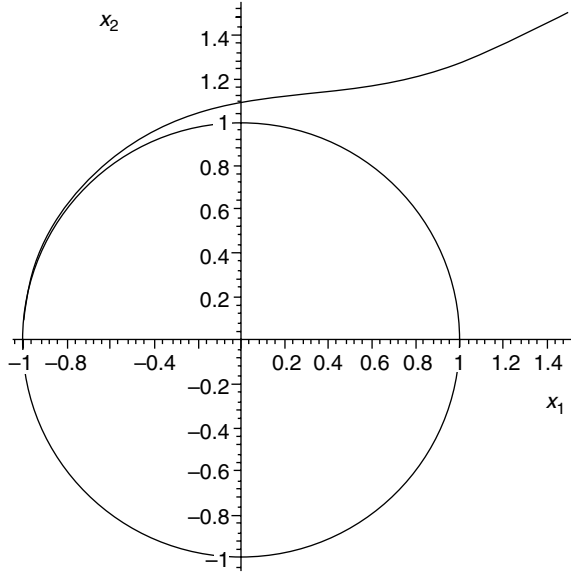
$$\frac{dx}{dt} = \begin{pmatrix} 1 - x_1^2 - x_2^2 & -1 \\ 1 & 1 - x_1^2 - x_2^2 \end{pmatrix} x, \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^2 \quad (2.13)$$

the fundamental elements of the phase portrait (see Figure 2.1) are the origin and the unit circle $\mathbf{C}_1$: starting from any initial condition out of the origin orbits converge to $\mathbf{C}_1$; otherwise, the state remains at the origin.

2.2.1.2 Existence

The CP may sometimes not have a solution or sometimes have many solutions. Indeed, the system

$$\begin{aligned} \frac{dx}{dt} &= |x|^{1/2}, \quad x \in \mathbb{R} \\ x(0) &= 0 \end{aligned} \quad (2.14)$$

**FIGURE 2.1**

Unit circle: simulation of (2.13).

has an infinite number of solutions (see Figure 2.2) defined by:

$$\varepsilon \in \mathbb{R}_+, \quad \phi_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$$

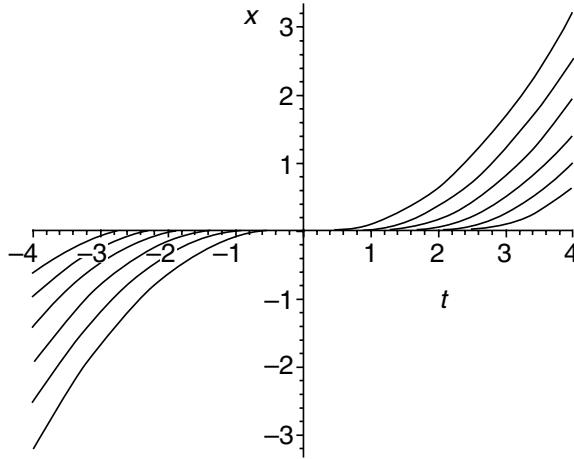
$$t \mapsto \phi_\varepsilon(t) = \begin{cases} 0 & \text{if } t_0 - \varepsilon \leq t \leq t_0 + \varepsilon \\ \frac{(t - t_0 - \varepsilon)^2}{4} & \text{if } t_0 + \varepsilon \leq t \\ -\frac{(t - t_0 + \varepsilon)^2}{4} & \text{if } t \leq t_0 - \varepsilon \end{cases} \quad (2.15)$$

Thus, one may wonder whether there exist conditions ensuring the existence of one or many solutions to the CP.

According to the smoothness of function f one can distinguish the following five cases A, B, C, D, and E.

CASE A

If function f is continuous with respect to x and eventually discontinuous with respect to t (but measurable), then there exist absolutely continuous solutions to the CP.

**FIGURE 2.2**

Infinite number of solutions to the CP of (2.14).

THEOREM 1 (Carathéodory, 1918) [8]

Assume that:

A1. f is defined for almost all t on a "barrel":

$$\mathcal{B} = \{(t, x) \in \mathcal{I} \times \mathcal{X} : |t - t_0| \leq a, \|x - x_0\| \leq b\} \quad (2.16)$$

A2. f is measurable with respect to t for all fixed x , continuous with respect to x for all fixed t and such that $\|f(t, x)\| \leq m(t)$ holds on $\mathcal{B}$, with m being a positive function which is Lebesgue-integrable on $|t - t_0| \leq a$.

Then, there exist at least one solution (absolutely continuous) to the CP which is defined at least on an interval like $[t_0 - \alpha, t_0 + \alpha]$, $\alpha \leq a$.

One can prove the existence of two solutions such that any other solution lies between these two [8, 18].

CASE B

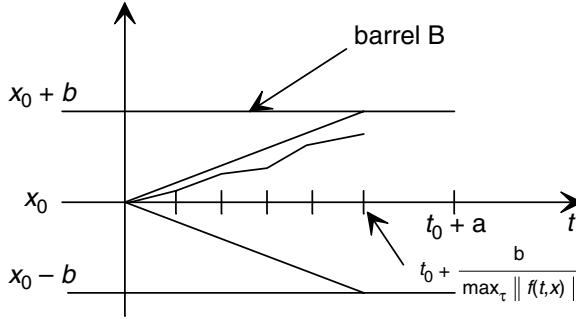
If the function f is continuous with respect to (t, x) , then there exist continuously differentiable solution (this is class $\mathcal{C}^1$ solution).

THEOREM 2 (Peano, 1886) [7]

Assume that:

B1. f is defined for all t on the "barrel" $\mathcal{B}$ defined by (2.16)

B2. f is continuous on $\mathcal{B}$ defined by (2.16)

**FIGURE 2.3**

Euler approximates.

Then there exist at least one solution to the CP belonging to the C^1 class of functions and defined at least on an interval like $[t_0 - \alpha, t_0 + \alpha]$, $\alpha = \min(a, b/\max_B \|f(t, x)\|)$.

The proof is based on *Euler approximates*, which are polygonal lines (see Figure 2.3) defined by:

$$\begin{cases} \phi_0 = x_0 \\ \phi_n(t) = \phi_n(t_{i-1}) + f(t_{i-1}, \phi_n(t_{i-1}))(t - t_{i-1}), & t_{i-1} < t \leq t_i \\ t_i = t_0 + \frac{i}{n}\alpha, & i = \{0, \dots, n\} \end{cases}$$

These functions constitute a family of equicontinuous functions defined on $[t_0 - \alpha, t_0 + \alpha]$, converging. Then, using the Ascoli–Arzela lemma one can extract a family ϕ'_n uniformly converging to a continuous function ϕ which satisfies:

$$\begin{aligned} \phi(t) &= \lim_{n \rightarrow +\infty} \phi'_n(t) = x_0 + \int_{t_0}^t \lim_{n \rightarrow +\infty} f(v, \phi'_n(v)) \, dv \\ &+ \lim_{n \rightarrow +\infty} \int_{t_0}^t \frac{d\phi'_n}{dt}(v) - f(v, \phi'_n(v)) \, dv. \end{aligned}$$

So ϕ is a solution of (2.11) since $\lim_{n \rightarrow +\infty} (d\phi'_n/dt)(v) - f(v, \phi'_n(v)) = 0$.

2.2.1.3 Extension, Uniqueness, and Global Solution

Obviously, in example (2.14), solutions to the CP exist ($f : x \mapsto \sqrt{|x|}$ is continuous) but are nonunique. To ensure uniqueness, the function f should

be “smoother than continuous”: for example, locally Lipschitz with respect to the second variable x is sufficient, as defined subsequently.

DEFINITION 3 f is said to be locally Lipschitz on $\mathcal{X}$ if: $\forall x_0 \in \mathcal{X}, \exists \delta > 0$ and $k(t)$ integrable:

$$\forall (x, y) \in \mathcal{B}_\delta(x_0) = \{x : \|x - x_0\| \leq \delta\} \Rightarrow \|f(t, x) - f(t, y)\| \leq k(t)\|x - y\|.$$

f is said globally Lipschitz on $\mathcal{X}$ if:

$$\exists k(t) \text{ integrable} : \forall (x, y) \in \mathcal{X}^2, \|f(t, x) - f(t, y)\| \leq k(t)\|x - y\|.$$

These properties are said to be *uniform* if k does not depends on t .

PROPOSITION 1

Any $\mathcal{C}^1(\mathcal{I} \times \mathcal{X})$ function with norm of the jacobian bounded by an integrable function, is locally Lipschitz. If, in addition, $\mathcal{X}$ is compact (this is to say closed and bounded since $\mathcal{X}$ is a subset of $\mathbb{R}^n$), then the function is globally Lipschitz.

Under assumption f being locally Lipschitz with respect to x , it may happen that a solution ϕ defined on $\mathcal{I}_1$ can be extended to a larger interval $\mathcal{I}_2 \supset \mathcal{I}_1$, and thus defines a new function $\tilde{\phi}$ defined on $\mathcal{I}_2 \supset \mathcal{I}_1$ and such that $\tilde{\phi} \mid \mathcal{I}_1 = \phi$. Thus, in order to not weigh down the notations, $\mathcal{I}(t_0, x_0) =]\alpha(t_0, x_0), \omega(t_0, x_0)[$ will indicate thereafter the greatest interval on which one can define a solution passing at time t_0 through x_0 and which cannot be extended: the solution will be known as *maximum solution*.

CASE C

If the function f is locally Lipschitz with respect to x and possibly discontinuous in t (but measurable), then there exist a unique maximum solution which is absolutely continuous.

THEOREM 3

If in Theorem 1, the assumption on the continuity given in A2 is replaced by “ f locally lipschitz on $\|x - x_0\| \leq b$,” then there exists a unique solution (absolutely continuous) to the problem of Cauchy defined on $\mathcal{I}(t_0, x_0) \supset \{t \in \mathcal{I} : |t - t_0| \leq \alpha\}$. Similarly, if f is continuous in (t, x) and locally Lipschitz in x , then there is a unique $\mathcal{C}^1$ solution to the CP.

The evidence of these results is based on the *Picard–Lindelöf* approximates:

$$\begin{cases} \phi_0 = x_0 \\ \phi_{n+1}(t) = x_0 + \int_{t_0}^t f(v, \phi_n(v)) \, dv \\ t \in [t_0 - \alpha, t_0 + \alpha], \quad \alpha = \min \left(a, \frac{b}{\max_{\mathcal{B}} \|f(t, x)\|} \right) \end{cases}$$

which shows that they converge uniformly towards a solution. Then, the uniqueness of the solution is shown by contradiction.

CASE D

If the function f has a norm which is bounded by a affine function, that is, $\forall (t, x) \in (\mathcal{I} \times \mathcal{X})$ (possibly almost everywhere): $\|f(t, x)\| \leq c_1 \|x\| + c_2$ with c_1 and c_2 strictly positive, then by using the lemma of Gronwall, one can conclude that any solution to the CP is defined on $\mathcal{I}$.

CASE E

If the system is “dissipative” and if f is locally lipschitz, then the CP admits a unique maximum solution for any $t \geq t_0$ ($\mathcal{I} = \mathbb{R}$, $\mathcal{X} = \mathbb{R}^n$). The property of dissipativity can be expressed such as: “there exists $\alpha \geq 0$, $\beta \geq 0$, $v \in \mathbb{R}^n$ such that for any $t \in \mathbb{R}$ and any $x \in \mathbb{R}^n$: $\langle x - v, f(t, x) \rangle \leq \alpha - \beta \|x\|^2$ ” or using Liapunov⁶ functions [17], such as “there exist V and $W : \mathbb{R}^n \mapsto \mathbb{R}_+$ continuous, positive definite on a compact $\mathcal{A}$ (i.e., $V(x) \geq 0$ and $V(x) = 0 \Leftrightarrow x \in \mathcal{A}$), such that for any $t \in \mathbb{R}$ and any $x \in \mathbb{R}^n \setminus \mathcal{A}$: $\langle \frac{\partial V}{\partial x}, f(t, x) \rangle \leq -W(x)$.”

2.2.1.4 Dependence of the Initial Conditions

THEOREM 4

Under assumptions of Theorem 3, the solution to the problem of Cauchy $t \mapsto \phi(t; t_0, x_0)$ defined on $\mathcal{I}(t_0, x_0)$ is continuous with respect to each one of its arguments.

In particular, if t is sufficiently close to t_0 , then $\phi(t; t_0, x_0)$ is also close to x_0 . This proximity can be studied for very large moments: it is the question of stability (see Section 2.3.2).

2.2.2 Classification

DEFINITION 4 The ODE (2.10) is said to be *autonomous* if the time variable t does not appear explicitly in the equation, thus (2.10) is in the form:

$$\frac{dx}{dt} = g(x), \quad t \in \mathcal{I}, \quad x \in \mathcal{X}$$

On the contrary (2.10) is said to be *non-autonomous*.

⁶Alexander Mikhaïlovich Liapunov, Russian mathematician and physicist. After completing his studies at the University of St. Petersburg (where he was the student of P.L. Tchebychev), he was an Assistant Professor and then Professor at the University of Kharkov. In 1902, he got a Professor position at the University of St. Petersburg.

⁷Notation $A \setminus B$ is the difference of two sets A and B : $A \setminus B = \{x \in A : x \notin B\}$.

If one knows all the solutions of an autonomous ODE which take the value x at time t , then one can get all the solutions which pass through this point at an other time just by using a time shift of the first set. Thus an autonomous ODE can only be used to model physical phenomena that do not depend on the initial time (e.g., a “rolling stone”). Note that the length of $\mathcal{I}(t_0, x_0)$ does not depend on the initial time.

DEFINITION 5 A nonlinear non-autonomous vector field $f(t, x)$ is said to be *T-periodic* if there exists a real number $T > 0$ such that for any t and any $x : f(t + T, x) = f(t, x)$.

If one knows the solutions on a time interval of length T , then one can get the solution on the entire time interval of definition just by time translation.

2.2.2.1 Autonomous Linear Case

When (2.10) is of the form:

$$\frac{dx}{dt} = Ax + b$$

the ODE is said to be *autonomous linear*. Then there exists a unique solution to the CP (since $Ax + b$ is globally uniformly Lipschitz) given by:

$$x(t) = e^{A(t-t_0)}x_0 + e^{At} \left(\int_{t_0}^t e^{-Av} dv \right) b$$

or $x(t) = \sum_{i=1}^r e^{\lambda_i t} p_i(t) + c$, where λ_i is the eigenvalue of A and $p_i(t)$ is the polynomial vector of degree less than the multiplicity order of the corresponding eigenvalues λ_i . Section 2.4 gives more precise results when $b = 0$.

This kind of model has the following properties:

1. Initial time has no influence on the time evolution of the state vector (the ODE is autonomous)
2. If $b = 0$ (resp. $b \neq 0$), then a linear combination (resp. convex) of the solution is still a solution: this is the *linear property* of the system.

For such a model one can note that after an infinite time the state vector x :

1. Either converges to a constant vector value (called an “equilibrium point”)
2. Or diverges (the norm of x becomes infinite)
3. Or oscillates: when one observes their evolutions, they evolve/move on a closed curve (as the circle): this is what is called a closed cycle (e.g., business cycle, cyclic population, and mass attached to a spring).

Lastly, if A and b depend on time, the system is known to be *linear non-autonomous* (or *nonstationary*): in addition to the aforementioned behaviors one finds the dependence of the solutions on the initial time. Note that when $A(t)$ is a T -periodic function continuous on $\mathbb{R}$, one can formally study the solutions using the theory of Floquet [13] which states that there is $P(t)$ a one-to-one transformation T -periodic and continuous, $z(t) = P(t)x(t)$, such that $\dot{z} = Mz + c(t)$, with M a constant matrix satisfying $M = \dot{P}(t) + P(t)A(t)P(t)^{-1}$ and $c(t) = P(t)b(t)$.

2.2.2.2 Nonlinear Autonomous Case

When (2.10) is of the form:

$$\frac{dx}{dt} = g(x) \quad (2.17)$$

the ODE is said to be *nonlinear autonomous*. Generally, one cannot get an explicit solution of these ODEs except for very particular cases. In addition to the aforementioned behaviors in the autonomous linear case, one can mention:

1. *Limit cycle*: they are closed curves in $\mathcal{X}$ toward or from which the trajectories of the system move.
2. *Phenomenon of chaos*: these behaviors, governed by ODEs (deterministic), are seemingly random. One of the characteristics is the sensitivity to the initial conditions: two very close initial conditions will give rise to two completely different evolutions (see Section 2.6).
3. *Strange attractor*: it is in general a set of noninteger dimensions, which expresses some “roughness” of the object. For example, a surface is of dimension 2, a volume is of dimension 3, whereas a snowflake having infinite ramifications is of noninteger dimension ranging between 2 and 3. When the trajectories move toward (resp. move away from) this set, it is known as “strange attractor (resp. repeller).” Often,

the presence of attractor or strange repeller is a sign of chaos, however, in certain cases the chaotic phenomenon is only transitory and disappears after a sufficiently long time (see Section 2.6.3).

2.2.3 Flow

Here, one supposes the existence and uniqueness of a solution (maximum) to the CP associated with (2.17), denoted as $\phi(t; t_0, x_0)$. If the vector field is *complete*, that is, $\mathcal{I}(x_0) = \mathbb{R}$, and if one knows a solution for a couple (t_0, x_0) , then one will have all the others (for fixed x_0) by time shifting. Consider the mapping which, for any initial condition, associates its maximum solution at the time t :

$$\begin{aligned}\Phi_g^t : \mathcal{X} &\rightarrow \mathcal{X} \\ x_0 &\mapsto \phi(t; 0, x_0)\end{aligned}$$

DEFINITION 6 If the vector field g of the ODE (2.17) makes it possible to generate a unique maximum solution for all (t_0, x_0) of $\mathbb{R} \times \mathcal{X}$ and defined on $\mathcal{I}(x_0) = \mathbb{R}$ (resp. on $[\alpha, \infty[$, on $[\alpha, \omega]$ with α and ω finite), then the generating application Φ_g^t is called a *flow* (resp. a *semi-flow* or a *local flow*).

According to the assumptions, Φ_g^t is one-to-one; therefore, there is at least a local flow. The justification of the notation Φ_g^t becomes obvious when computing the flow of an homogeneous autonomous linear ODE: $\dot{x} = Ax$, $\Phi_g^t = e^{At}$. If g is of class $\mathcal{C}^k$ (resp. $\mathcal{C}^\infty$, analytic), the associated flow Φ_g^t , is a local diffeomorphism of class $\mathcal{C}^k$ (resp. $\mathcal{C}^\infty$, analytic) for any time t where it is defined. In particular, if the flow Φ_g^t is defined for any $t \in \mathbb{R}$, then it defines a one parameter group of local diffeomorphisms of class $\mathcal{C}^k$ (resp. $\mathcal{C}^\infty$, analytic) (see p. 55–63 of [2]):

$$\Phi_g^t : x_0 \mapsto \Phi_g^t(x_0) \text{ is } \mathcal{C}^\infty \quad (2.18)$$

$$\Phi_g^t \circ \Phi_g^s = \Phi_g^{t+s} \quad (2.19)$$

$$\Phi_g^0 = Id \quad (2.20)$$

One deduces, $\forall t \in \mathbb{R}, \forall x_0 \in \mathcal{X}$:

$$\Phi_g^t(x_0) = \Phi_{-g}^{-t}(x_0) \quad (2.21)$$

$$\Phi_g^t \circ \Phi_g^{-t} = \Phi_g^{t-t} = Id \quad (2.22)$$

$$(\Phi_g^t)^{-1} = \Phi_g^{-t} = \Phi_{-g}^t \quad (2.23)$$

The duality characterized by (2.23) is important since, if one knows the phase portrait of the dynamic system (2.17) for positive times, its dual for negative times is obtained quite simply by reversing the go through direction of the orbits. This property is used in the trajectory-reversing method allowing, in two dimension (and sometimes three), to precisely determine the majority of the phase portraits by combining the qualitative study of the nonlinear vector field to simulations [5, 6, 9, 10, 21].

The *Lie bracket* (or commutator) defined by:

$$[g_1, g_2] = \left(\frac{\partial g_2}{\partial x} g_1 - \frac{\partial g_1}{\partial x} g_2 \right)$$

gives the condition of commutation of two vector flows $\Phi_{g_1}^t$ and $\Phi_{g_2}^s$.

THEOREM 5

Let g_1 and g_2 be two C^∞ complete vector fields defined on $\mathcal{X}$ (e.g., $\mathbb{R}^n$). Then:

$$\forall t, \forall s, \quad \Phi_{g_1}^t \circ \Phi_{g_2}^s = \Phi_{g_2}^s \circ \Phi_{g_1}^t \iff [g_1, g_2] = 0$$

In automatic control, the noncommutation of the vector fields has a very important application since it makes it possible to characterize the atteignability (local version of the controllability) of a controlled system of type $\dot{x} = g_1(x) + g_2(x)u$ [16].

2.3 Some Characterizations of Behaviors

Recall the ODE considered (2.10):

$$\frac{dx}{dt} = f(t, x), \quad t \in \mathcal{I}, \quad x \in \mathcal{X}$$

2.3.1 Remarkable Sets

2.3.1.1 Equilibrium Point

For some initial conditions, the system remains “frozen,” that is, the state does not evolve/move any more: one will speak then about equilibrium points.

DEFINITION 7 $x_e \in \mathcal{X}$ is an equilibrium point for the system (2.17) if all the solutions $\phi(t; 0, x_e)$ of (2.17) are defined on $[0, +\infty[$ and satisfy:

$$\phi(t; 0, x_e) = x_e, \quad \forall t \in [0, +\infty[$$

Example 2

The solution to CP (2.1), $x(0) = x_0$ is given by (2.12). It is easy to check that $x = 0$ and $x = x_{\max}$ are equilibrium points.

One can give a similar definition for (2.10) by taking into account the fact that the solutions then depend on the initial time. Thus, a point can be an equilibrium only for certain initial times. If x_e is an equilibrium point, then for the system to remain at this point it is necessary that the speed is null (i.e., $g(x_e) = 0$). However, this condition alone is not sufficient as shown with the study of (2.14): there are $x_e = 0$ solutions of $\sqrt{x} = 0$, but there is an infinite number of solutions which leave this point (see Section 2.2.1).

THEOREM 6 ([11])

x_e is an equilibrium point for the system (2.17) if and only if:

1. (2.17) admits a unique solution defined on $[t_0, +\infty[$ to the Cauchy problem
2. $g(x_e) = 0$

Thereafter, one will consider that the equilibrium point is the origin: indeed, the study of (2.17) in the neighborhood of an equilibrium point x_e is brought back, by the change of coordinates $y = x - x_e$, to the study of $\dot{y} = g(y + x_e)$, having for equilibrium ($y = 0$).

Example 3

The Volterra–Lotka system is a simple model of fight between two species. In 1917, during the war, the biologist Umberto d’Ancona noted an increase in the number of selacians (sharks) in the northern part of the Adriatic Sea. In order to explain this phenomenon, he called upon his father-in-law, the mathematician Vito Volterra, who explained this phenomenon in the following way. Let an infinite volume of water (e.g., Adriatic Sea) be populated by two species: one, carnivorous (C: selacians), chasing the other, herbivorous (H: shrimps). Let us note x and y the respective numbers of individuals of the species (H) and (C). If only the species (H) populated the sea, it would develop with an exponential rate⁸ and the speed of growth of the species (H) would be: $(dx/dt) = \alpha x$, with $\alpha > 0$. On the other hand, the development and survival of species (C), if alone, cannot be assured. Therefore, its speed of variation would be: $(dy/dt) = -\beta y$, with $\beta > 0$. When the two species cohabit, the carnivorous (C) devour the herbivores (H). By assuming that with each meeting of a carnivore with an herbivore, the latter is devoured and that the number of meetings is proportional to the product of the volumic densities of the two species (thus, also with xy), one can conclude that the evolution of the two species is governed by the

⁸One makes the assumption here that its development is limited neither by space nor by the quantity of food.

differential connection:

$$\begin{cases} \frac{dx}{dt} = \alpha x - \gamma xy & (\text{herbivorous}) \\ \frac{dy}{dt} = -\beta y + \delta xy & (\text{carnivorous}) \end{cases} \quad (2.24)$$

with $\alpha, \beta, \gamma, \delta$ being positive numbers. In this case, the state variables are introduced in a natural way: x, y . One can suppose a priori that the state space is the quarter of plan $\mathbb{R}_+^2$. Theorem 3 makes it possible to guarantee the existence and uniqueness of the solutions, and Theorem 6, the existence of two equilibrium points $(0,0)$ and $(\beta/\delta, \alpha/\gamma)$. By separating the variables according to $dx/(x(\alpha - \gamma y)) = dy/(y(-\beta + \delta x))$, one can show that $H(x, y) = [\alpha \ln(y) - \gamma y] + [\beta \ln(x) - \delta x]$ is a constant function along the solutions of (2.24). One shows thus that, for any initial condition strictly included in the quarter of strictly positive plan, the orbits of the system are closed. Moreover, solutions are defined on $\mathbb{R}$: one obtains a flow whose phase portrait is given in Figure 2.4 (simulation for $\alpha = \beta = \gamma = \delta = 1$).

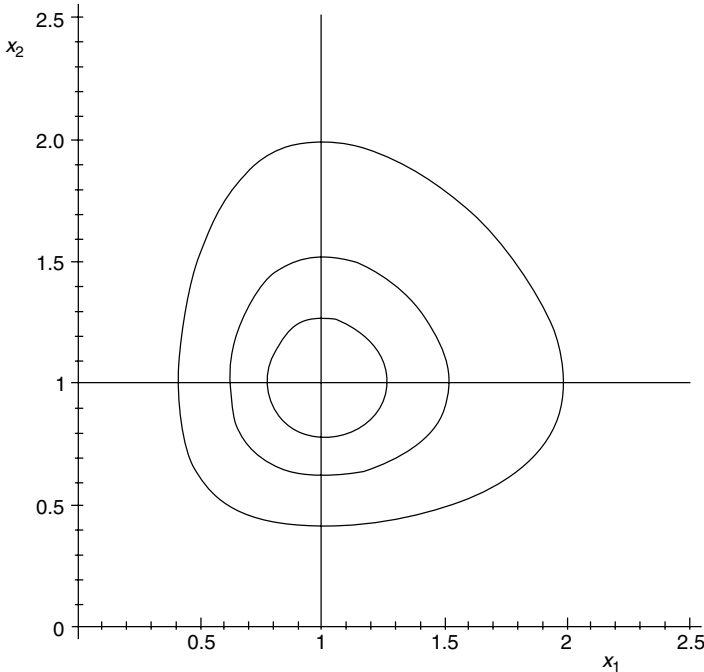


FIGURE 2.4
Limit cycle of (2.24).

The orbits are centered around the equilibrium point $(\beta/\delta, \alpha/\gamma)$. Before the war, the activity of fishing was more important (one takes into account fishing activity “ $-q_x x$ ” and “ $-q_y y$ ” in (2.24), with q_x, q_y positive): that is, the couple of parameters $(\alpha, -\beta)$ are replaced by $(\alpha - q_x, -\beta - q_y)$; therefore, the equilibrium point $(\beta/\delta, \alpha/\gamma)$ is replaced by $((\beta + q_y)/\delta, (\alpha - q_x)/\gamma)$. This explains a displacement of the cycle during the war and therefore an increase in the number of celacians.

DEFINITION 8 One classifies the equilibrium points (x_e) of (2.17) in two categories:

1. *Hyperbolic points* (or nondegenerated): it is those for which the corresponding jacobian matrix⁹ $J_g(x_e)$ does not have any eigenvalue with null real part ($J_g(x_e)$ is also known as *hyperbolic*).
2. *Nonhyperbolic points* (or *degenerated*): it is those for which the jacobian matrix $J_g(x_e)$ has at least one eigenvalue with null real part ($J_g(x_e)$ is known as *degenerated*).

As we will see in Section 2.5.2, this distinction is important since, for any hyperbolic point, one knows the behavior of the solutions locally, whereas it is not inevitably the case for the nonhyperbolic points.

2.3.1.2 Orbits: Periodic, Closed, Homoclinic, and Heteroclinic

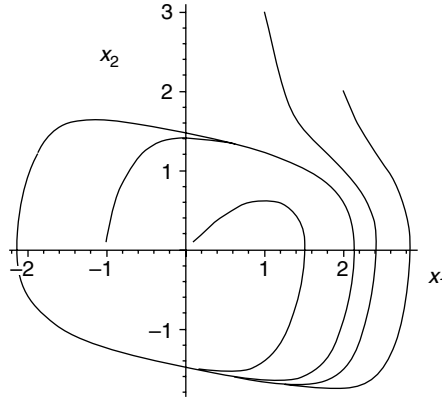
The study of nonlinear systems highlights particular orbits:

1. The *closed orbits* which are an extension of the fixed points (or equilibrium) since, if one lets a system to evolve starting from an initial condition belonging to this orbit, then it will continue to evolve on this orbit;
2. Homoclinic and heteroclinic orbits which connect equilibrium points.

The following definitions are inspired by p. 87–88 of [24], p. 8 of [25], p. 113–117 of [14], and [12].

DEFINITION 9 The solution $\phi(t; t_0, x_0)$ is *T-periodic* (periodic with period T), if $\mathcal{I}(t_0, x_0) = \mathbb{R}$ and if there exists a positive real λ , such that for any real t one has $\phi(t + \lambda; t_0, x_0) = \phi(t; t_0, x_0)$. The smallest positive number λ noted T is called the period of the solution. In this case, the corresponding orbit is a *periodic orbit* of period T (or *T-periodic orbit*).

⁹If g is a vector field on $\mathbb{R}^n$, then its jacobian matrix at the point x is the matrix $((\partial g_i / \partial x_j)(x))$.

**FIGURE 2.5**

Periodic closed orbit of the Van der Pol oscillator (2.25).

DEFINITION 10 γ is a closed orbit if γ is an orbit and a Jordan curve (i.e., homeomorphic¹⁰ to a circle).

Any orbit corresponding to a nontrivial T -periodic solution (nonidentical to an equilibrium point) is a closed orbit. The reciprocal one is true in the case of autonomous systems.

Example 4

If one looks again at the Van der Pol equation, that is, the Equation (2.3) with $f(x) = (x^3 - 2\mu x)$, then denoting by $i_L = -x_2$, $v_C = x_1$, $L = C = 1$, (2.3) becomes:

$$\begin{aligned}\frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= 2\mu x_2 - x_2^3 - x_1\end{aligned}\tag{2.25}$$

Thus, for $\mu > 0$, one can show (p. 211–227 of [15]) the existence of γ , a periodic orbit represented in Figure 2.5.

DEFINITION 11 A *homoclinic orbit* is an orbit which connects an equilibrium point to itself. If an orbit connects two distinct equilibrium points it is known as *heteroclinic* (Figure 2.6 and Figure 2.7).

¹⁰A homeomorphism is a continuous morphism. Thus, the Jordan curve is a curved obtain by continuous transformation starting from a circle.

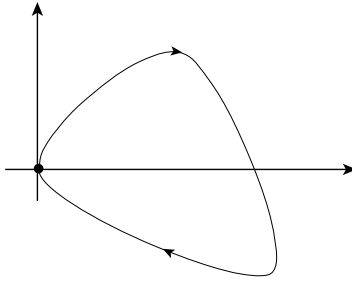


FIGURE 2.6
Homoclinic orbit.

2.3.2 Properties

In this section, one considers the system (2.10) and assumes that there is at least one solution to the CP.

2.3.2.1 Invariance

Physical systems often have the tendency, in certain configurations, to satisfy a principle of less effort: “here I am, here I remain” (equilibrium points, periodic orbits, etc.). This property of invariance can be extended to more complex sets.

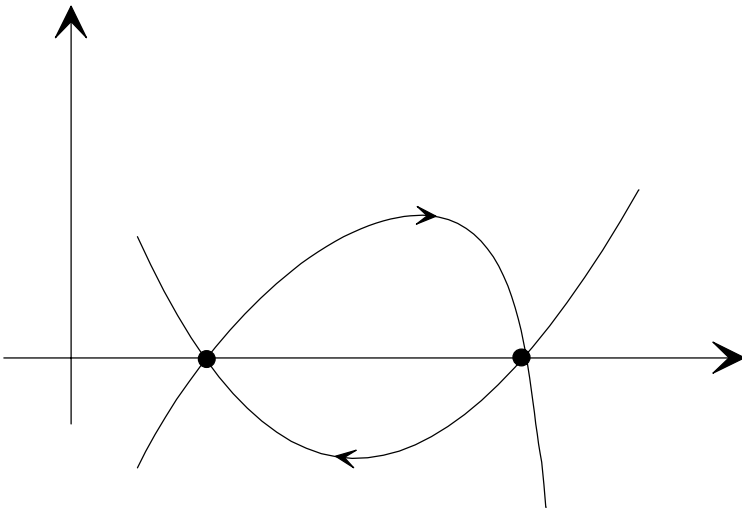


FIGURE 2.7
Heteroclinic orbit.

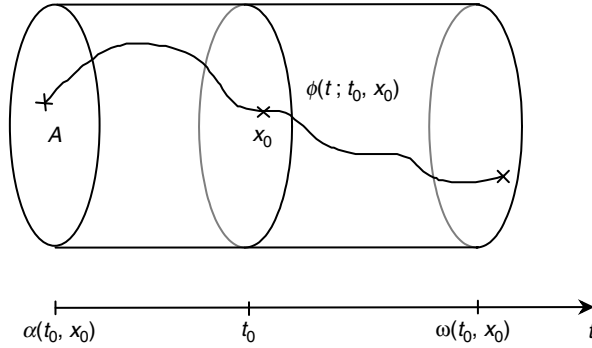


FIGURE 2.8
Invariance of $\mathcal{A}$.

DEFINITION 12 Let $\mathcal{J} \subset \mathcal{I}$. A nonempty compact set $\mathcal{A} \subset \mathcal{X}$ is $\mathcal{J}$ -invariant if: $\forall t_0 \in \mathcal{J}, \forall x_0 \in \mathcal{A}, \forall t \in \mathcal{I}(t_0, x_0) : \phi(t; t_0, x_0) \in \mathcal{A}$ (Figure 2.8).

2.3.2.2 Liapunov Stability

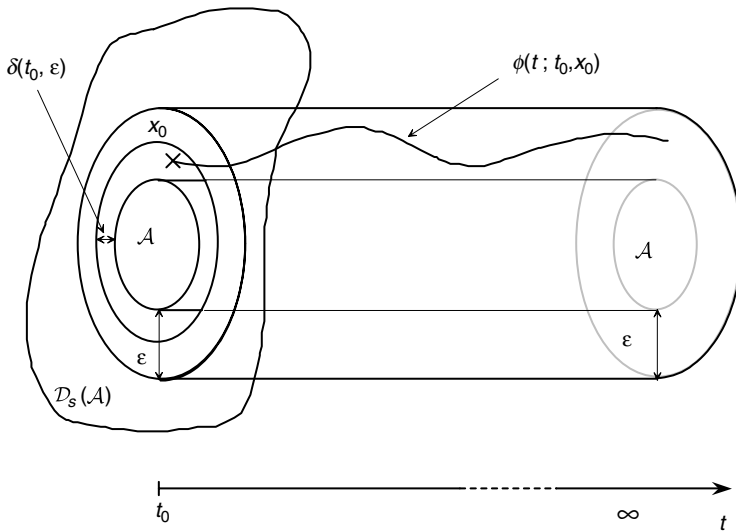
Remarkable sets (equilibrium points, periodic orbits, etc.) can characterize configurations with minimal energy for a physical system. These systems can tend to seek one of these configurations rather than another, according to the concepts of stability. For example, the pendulum with mass in vertical position (2.7) has two equilibria: one above the horizontal, $\theta = \pi, \dot{\theta} = 0$, the other below $\theta = 0, \dot{\theta} = 0$. It is well known that the mass naturally tends to the bottom position rather than to the upper one. The bottom position of the equilibrium is stable, the other one unstable. From another point of view, the maximum solution $x(t; t_0, x_0)$ of an ODE is continuous with respect to the three variables t, t_0, x_0 (under some conditions, see Section 2.2.1). Therefore, if two solutions $x(t; t_0, x_{01})$ and $x(t; t_0, x_{02})$ are taken with x_{01} close to x_{02} , continuity implies that these two solutions are close on some time interval $[t_0, t]$, without any indication on the size of this interval. One can obtain a proximity of these two solutions on an interval of rather large time as stated in the problem of Liapunov stability for a particular solution (equilibrium point, periodic orbit, set or a given trajectory).

Thereafter, $\mathcal{A}$ is a non-empty compact set (e.g., an equilibrium point) of $\mathcal{X}$ endowed with a distance d . $\rho(x, \mathcal{A}) = \inf_{y \in \mathcal{A}} d(x, y)$ is a distance from the point x to the set $\mathcal{A}$. Lastly, $\mathcal{I}(t_0, x_0) =]\alpha(t_0, x_0), +\infty[$.

DEFINITION 13 $\mathcal{A}$ is Liapunov stable with respect to $\mathcal{J} \subset \mathcal{I}$ for (2.10) if:

$\forall t_0 \in \mathcal{J}, \forall \varepsilon > 0, \exists \delta(t_0, \varepsilon) > 0$ such that:

$$\forall x_0 \in \mathcal{X} : \rho(x_0, \mathcal{A}) \leq \delta(t_0, \varepsilon) \Rightarrow \rho(\phi(t; t_0, x_0), \mathcal{A}) \leq \varepsilon, \quad \forall t \geq t_0$$

**FIGURE 2.9**

Stability of the $\mathcal{A}$ and its stability domain $\mathcal{D}_s(\mathcal{A})$.

When $\mathcal{J} = \mathcal{I} = \mathbb{R}$, $\mathcal{A}$ is said to be *Liapunov stable*. When $\delta(t_0, \epsilon) = \delta(\epsilon)$ do not depend on t_0 , the stability property is said to be *uniform*.

In the particular case of the autonomous nonlinear systems (2.17), for any neighborhood of $\mathcal{A}$, there is a positively invariant neighborhood of $\mathcal{A}$ (included in the first) (see p. 58 of [4]).

These definitions can be stated in more general topological terms: for example, an equilibrium point x_e is Liapunov stable for (2.17) if, for any neighborhood $\mathcal{V}(x_e)$ of x_e , there is a neighborhood $\mathcal{W}(x_e)$ of x_e such that: $x_0 \in \mathcal{W}(x_e) \Rightarrow \phi(t; t_0, x_0) \in \mathcal{V}(x_e), \forall t \geq t_0$. In general, for (2.10), this definition is stated as: for all $t_0 \in \mathcal{J}$ and any neighborhood $\mathcal{V}(\mathcal{A})$ of $\mathcal{A}$, there exists a neighborhood $\mathcal{W}(t_0, \mathcal{V})$ of $\mathcal{A}$ such that any trajectory resulting from this neighborhood $\mathcal{W}(t_0, \mathcal{V})$ at the time t_0 evolves in the first neighborhood $\mathcal{V}(\mathcal{A})$ without leaving it (see Figure 2.9). However, if t_0 is fixed, for each neighborhood $\mathcal{V}(\mathcal{A})$, it is useful to know the largest of these neighborhood $\mathcal{W}(t_0, \mathcal{V})$, which will be denoted $\mathcal{D}_s(t_0, \mathcal{V}, \mathcal{A})$: this corresponds to the concept of *stability domain* of stability [the intersection of the sets $\mathcal{D}_s(t_0, \mathcal{V}, \mathcal{A})$].

DEFINITION 14 For (2.10), $\mathcal{D}_s(t_0, \mathcal{A})$ is the *Liapunov stability domain* with respect to t_0 of the set $\mathcal{A}$ if:

1. $\forall \epsilon > 0, \mathcal{D}_s(t_0, \epsilon, \mathcal{A}) \subset \mathcal{X}$ is a neighborhood of $\mathcal{A}$

2. For $\varepsilon > 0$, $x_0 \in \mathcal{D}_s(t_0, \varepsilon, \mathcal{A})$ iff: $\rho(\phi(t; t_0, x_0), \mathcal{A}) \leq \varepsilon, \forall t \geq t_0$
3. $\mathcal{D}_s(t_0, \mathcal{A}) = \cup_{\varepsilon > 0} \mathcal{D}_s(t_0, \varepsilon, \mathcal{A})$

$\mathcal{D}_s(\mathcal{J}, \mathcal{A})$ is the Liapunov stability domain with respect to $\mathcal{J}$ of the set $\mathcal{A}$ if:

1. $\forall t_0 \in \mathcal{J}$, $\mathcal{D}_s(t_0, \mathcal{A})$ exists
2. $\mathcal{D}_s(\mathcal{J}, \mathcal{A}) = \cup_{t_0 \in \mathcal{J}} \mathcal{D}_s(t_0, \mathcal{A})$

$\mathcal{D}_s(\mathcal{A})$ is the Liapunov stability domain of $\mathcal{A}$ if: $\mathcal{D}_s(\mathcal{A}) = \mathcal{D}_s(\mathcal{J} = \mathbb{R}, \mathcal{A})$.

DEFINITION 15 For (2.10), $\mathcal{D}_{us}(\mathcal{J}, \mathcal{A})$ is the *Liapunov uniform stability domain* with respect to $\mathcal{J}$ of the set $\mathcal{A}$ if:

1. $\mathcal{D}_{us}(\mathcal{J}, \mathcal{A})$ is a neighborhood of $\mathcal{A}$
2. $x_0 \in \mathcal{D}_{us}(\mathcal{J}, \mathcal{A})$ iff: $\forall t_0 \in \mathcal{J}, \forall \varepsilon > 0, \rho(\phi(t; t_0, x_0), \mathcal{A}) \leq \varepsilon, \forall t \geq t_0$

$\mathcal{D}_{us}(\mathcal{A})$ is the Liapunov uniform stability domain of $\mathcal{A}$ if: $\mathcal{D}_{us}(\mathcal{A}) = \mathcal{D}_{us}(\mathcal{J} = \mathbb{R}, \mathcal{A})$.

DEFINITION 16 Let us symbolize schematically by $(\bullet)$ one of the four following qualifiers: Liapunov stable with respect to $\mathcal{J}$, Liapunov stable, Liapunov uniformly stable with respect to $\mathcal{J}$, Liapunov uniformly stable. If $\mathcal{D}_{(\bullet)}(\mathcal{J}, \mathcal{A}) = \mathcal{X}$ (the state space), the $\mathcal{A}$ is *globally* $(\bullet)$, on the contrary $\mathcal{A}$ is *locally* $(\bullet)$.

REMARK 1

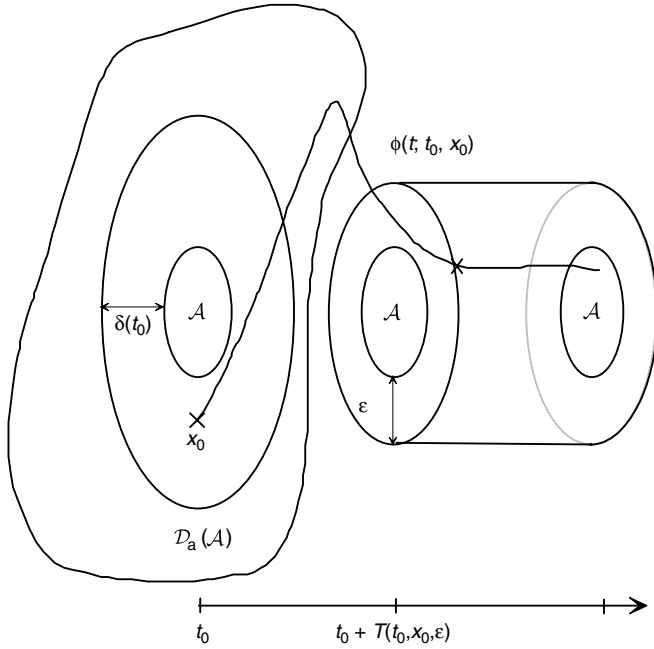
The definitions of stability (Definition 13) can be replaced by: the set $\mathcal{A}$ is $(\bullet)$ if $\mathcal{D}_{(\bullet)}$ is a non-empty domain.

REMARK 2

Once again, these definitions can be stated in terms of neighborhoods. The previous construction leads us to proceed as follows: if, for each $\mathcal{V}(\mathcal{A})$, one builds the union of $\mathcal{D}_s(t_0, \mathcal{V}(\mathcal{A}), \mathcal{A})$ (which is similar to the $\mathcal{D}_s(t_0, \varepsilon, \mathcal{A})$), then one obtains the Liapunov stability domain with respect to t_0 of $\mathcal{A}$ denoted: $\mathcal{D}_s(t_0, \mathcal{A}) = \cup_{\mathcal{V}(\mathcal{A})} \mathcal{D}_s(t_0, \varepsilon, \mathcal{A})$. If t_0 is the range with the interval $\mathcal{J}$, one obtains the Liapunov stability domain of $\mathcal{A}$ with respect to $\mathcal{J}$, denoted $\mathcal{D}_s(\mathcal{J}, \mathcal{A})$.

2.3.2.3 Attractivity

The attractivity property of a set means that solutions asymptotically tend to this set.

**FIGURE 2.10**

Attractivity of the set $\mathcal{A}$ and its attractivity domain $\mathcal{D}_a(\mathcal{A})$.

DEFINITION 17 $\mathcal{A}$ is attractive with respect to $\mathcal{J} \subset \mathcal{I}$ for (2.10) if: $\forall t_0 \in \mathcal{J}, \exists \delta(t_0) > 0$ such that $\forall x_0 \in \mathcal{X} : \rho(x_0, \mathcal{A}) \leq \delta(t_0) \Rightarrow \lim_{t \rightarrow \infty} \rho(\phi(t; t_0, x_0), \mathcal{A}) = 0$, i.e., $\forall \varepsilon > 0, \exists T(t_0, x_0, \varepsilon) > 0 : \forall t \geq t_0 + T(t_0, x_0, \varepsilon), \rho(\phi(t; t_0, x_0), \mathcal{A}) \leq \varepsilon$. When $\mathcal{J} = \mathcal{I} = \mathbb{R}$, $\mathcal{A}$ is called attractive. When $\delta(t_0) = \delta$ does not depend on t_0 and $T(t_0, \varepsilon, x_0) = T(\varepsilon)$ does not depend on t_0 and x_0 , then the attractivity property is said to be *uniform*.

This concept can be formulated in terms of neighborhoods. For example, for all $t_0 \in \mathcal{J}$ and any $\mathcal{V}(\mathcal{A})$ neighborhood of $\mathcal{A}$, there exists $\mathcal{W}(t_0, \mathcal{V})$ a neighborhood of $\mathcal{A}$ such that, for any trajectory resulting from this neighborhood $\mathcal{W}(t_0, \mathcal{V})$ at the moment t_0 , there is a time $T(t_0, x_0, \mathcal{V}) > 0$ such that the trajectory evolves in $\mathcal{W}(t_0, \mathcal{V})$ without leaving there from the moment $t_0 + T(t_0, x_0, \mathcal{V})$ (see Figure 2.10). However, for t_0 and $\mathcal{V}(\mathcal{A})$, a given neighborhood of $\mathcal{A}$, it is useful to know the largest of these neighborhoods $\mathcal{W}(t_0, \mathcal{V})$ which will note $\mathcal{D}_a(t_0, \mathcal{V}, \mathcal{A})$: this led to the concept of *attractivity domain*, intersection of $\mathcal{D}_a(t_0, \mathcal{V}, \mathcal{A})$ (with respect to the $\mathcal{V}$).

DEFINITION 18 For (2.10), $\mathcal{D}_a(t_0, \mathcal{A})$ is the *attractivity domain* of $\mathcal{A}$ with respect to t_0 if:

1. $\mathcal{D}_a(t_0, \mathcal{A}) \subset \mathcal{X}$ is a neighborhood of $\mathcal{A}$

2. For $\varepsilon > 0$, $x_0 \in \mathcal{D}_a(t_0, \varepsilon, \mathcal{A})$ if and only if: $\exists T(t_0, \varepsilon) > 0$ such that $\forall t \geq t_0 + T(t_0, \varepsilon)$, $\rho(\phi(t; t_0, x_0), \mathcal{A}) \leq \varepsilon$

$\mathcal{D}_a(\mathcal{J}, \mathcal{A})$ is the attractivity domain of $\mathcal{A}$ with respect to $\mathcal{J}$ if:

1. $\forall t_0 \in \mathcal{J}$, $\mathcal{D}_a(t_0, \mathcal{A})$ exists
2. $\mathcal{D}_a(\mathcal{J}, \mathcal{A}) = \cup_{t_0 \in \mathcal{J}} \mathcal{D}_a(t_0, \mathcal{A})$

$\mathcal{D}_a(\mathcal{A})$ is the attractivity domain of $\mathcal{A}$ if: $\mathcal{D}_a(\mathcal{A}) = \mathcal{D}_a(\mathcal{J} = \mathbb{R}, \mathcal{A})$.

DEFINITION 19 For (2.10), $\mathcal{D}_{ua}(\mathcal{J}, \mathcal{A})$ is the *uniform attractivity domain* of $\mathcal{A}$ with respect to $\mathcal{J}$ if:

1. $\mathcal{D}_{ua}(\mathcal{J}, \mathcal{A})$ is a neighborhood of $\mathcal{A}$
2. For $\varepsilon > 0$, $x_0 \in \mathcal{D}_{ua}(\mathcal{J}, \mathcal{A})$ if and only if: $\exists T(\varepsilon) > 0$, $\forall t_0 \in \mathcal{J}$, such that $\rho(\phi(t; t_0, x_0), \mathcal{A}) \leq \varepsilon$, $\forall t \geq t_0 + T(\varepsilon)$.

$\mathcal{D}_{au}(\mathcal{A})$ is the uniform attractivity domain of $\mathcal{A}$ if: $\mathcal{D}_{au}(\mathcal{A}) = \mathcal{D}_{ua}(\mathcal{J} = \mathbb{R}, \mathcal{A})$.

DEFINITION 20 Let us symbolize schematically by $(\bullet)$ one of the four following qualifiers: attractive with respect to $\mathcal{J}$, attractive, uniformly attractive with respect to $\mathcal{J}$, uniformly attractive. If $\mathcal{D}_{(\bullet)}(\mathcal{J}, \mathcal{A}) = \mathcal{X}$ (the state space), then $\mathcal{A}$ is globally $(\bullet)$; on the contrary, $\mathcal{A}$ is locally $(\bullet)$.

REMARK 3

The definitions of attractivity (Definition 17) can be replaced by: the set $\mathcal{A}$ is $(\bullet)$ if the domain $\mathcal{D}_{(\bullet)}$ is non-empty.

2.3.2.4 Asymptotic Stability

The attractivity property of a set expresses the convergence of the solutions into this set after an infinite time, and this, independently of possible excursions during the transient phase. Stability property expresses the proximity of solutions throughout the evolution, but without guaranteeing convergence. These two properties are thus distinct and complementary. Their combination corresponds to the concept of asymptotic stability.

DEFINITION 21 $\mathcal{A}$ is asymptotically stable with respect to $\mathcal{J} \subset \mathcal{I}$ if it is Liapunov stable $\mathcal{J}$ and attractive with respect to $\mathcal{J}$. When $\mathcal{J} = \mathcal{I} = \mathbb{R}$, $\mathcal{A}$ is known as *asymptotically stable*. If the properties of stability and attractivity are uniform, then the obtained asymptotic property is known as *uniform*. The various asymptotic stability domains associated to the aforementioned

properties are defined as being the intersections of the domains of stability and attractivity.

Note that for a nonlinear system, a set can be attractive without being stable, and vice versa. The following example illustrates this independence of the two properties.

Example 5

Consider the following ODE:

$$\begin{aligned}\frac{dx}{dt} &= x \left(1 - \sqrt{x^2 + y^2} \right) - \frac{y}{2} \left(1 - \frac{x}{\sqrt{x^2 + y^2}} \right) \\ \frac{dy}{dt} &= y \left(1 - \sqrt{x^2 + y^2} \right) + \frac{x}{2} \left(1 - \frac{x}{\sqrt{x^2 + y^2}} \right)\end{aligned}$$

The origin $(0,0)$ is an unstable equilibrium point and the equilibrium $(1,0)$ is attractive but unstable: the phase portrait is given in Figure 2.11.

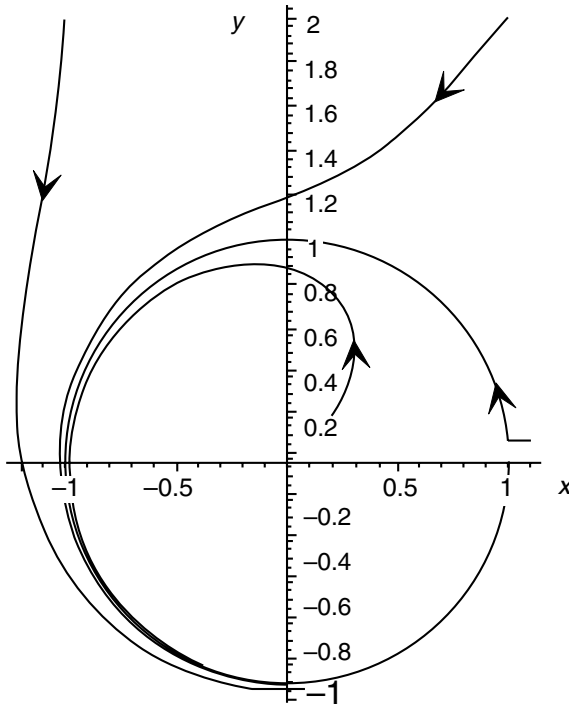


FIGURE 2.11

Equilibrium $(1,0)$ is attractive and unstable.

Example 6

The solution to the CP of (2.1) $x(0) = x_0$ is given by (2.12). It is thus easy to check that the equilibrium $x = 0$ is not attractive ($\lim_{t \rightarrow \infty} \phi(t; 0, x_0) = \lim_{t \rightarrow \infty} (x_0 x_{\max} / x_0 + e^{-ax_{\max} t} (x_{\max} - x_0)) = x_{\max}$) and that the equilibrium $x = x_{\max}$ is asymptotically stable. Indeed, it is attractive ($\lim_{t \rightarrow \infty} \phi(t; 0, x_0) = x_{\max}$) and stable since:

$$\phi(t; 0, x_0) - x_{\max} = \frac{x_{\max}(x_{\max} - x_0)e^{-ax_{\max} t}}{x_0 + (x_{\max} - x_0)e^{-ax_{\max} t}}$$

thus for $\varepsilon > 0$, if $|x_0 - x_{\max}| < \varepsilon$, $|\phi(t; 0, x_0) - x_{\max}| < x_{\max}|(x_{\max} - x_0)/(x_0 + (x_{\max} - x_0))| < \varepsilon$.

For this example, we could study asymptotic stability from the analytical expression of the solutions. However, for an ODE (2.10) for which, in general, one cannot get explicit solutions, it is important to have a criterion allowing to study the question of stability without having to calculate the solutions: these are the results of Section 2.4 (the first method of Liapunov) and others based on the use of Liapunov function (the second method of Liapunov [17, 23]).

2.3.2.5 Exponential Stability

The notion of exponential stability contains an additional information: the speed of convergence toward the set $\mathcal{A}$.

DEFINITION 22 $\mathcal{A}$ is exponentially stable with respect to $\mathcal{J} \subset \mathcal{I}$ if: $\forall t_0 \in \mathcal{J}, \exists \delta(t_0) > 0, \exists \alpha(\delta) > 0, \exists \beta(\delta) \geq 1$ such that $\forall x_0 \in \mathcal{X}, \rho(x_0, \mathcal{A}) \leq \delta(t_0)$ implies:

$$\rho(\phi(t; t_0, x_0), \mathcal{A}) \leq \beta \rho(x_0, \mathcal{A}) \exp(-\alpha(t - t_0)), \quad \forall t \geq t_0 \quad (2.26)$$

When $\mathcal{J} = \mathcal{I} = \mathbb{R}$, $\mathcal{A}$ is called *exponentially stable*. When $\delta(t_0) = \delta$ do not depend on t_0 and $\alpha(\delta) = \alpha, \beta(\delta) = \beta$ do not depend on δ , then the stability property is called *uniform*. α is then called the *rate of exponential convergence*.

Just as for the preceding concepts, one can define the corresponding exponential stability domains. The definition of these domains takes into account the obtained pair (α, β) : for example, $\mathcal{D}_{ue}(\mathcal{A}, \alpha, \beta)$ is the greatest neighborhood of $\mathcal{A}$ for which the overevaluation (2.26) holds.

2.4 Autonomous Linear Case

2.4.1 Formal Computation of Solution

Let us consider the following linear autonomous system:

$$\dot{x} = Ax, \quad x \in \mathbb{R}^n \quad (2.27)$$

The solution to the CP associated to (2.27), $x(t_0) = x_0$, is explicitly given by:

$$x(t) = \exp(A(t - t_0))x_0 \quad (2.28)$$

And for the following:

$$\dot{x} = Ax + b(t), \quad x \in \mathbb{R}^n \quad (2.29)$$

the solution is given by:

$$x(t) = \exp(A(t - t_0)) \left(x_0 + \int_{t_0}^t \exp(-A(v - t_0))b(v) \, dv \right) \quad (2.30)$$

which, when b is constant, is given by:

$$x(t) = \exp(A(t - t_0))x_0 + \left(\int_{t_0}^t \exp A(t - v) \, dv \right) b \quad (2.31)$$

which, when A is nonsingular, reads as:

$$x(t) = \exp(A(t - t_0))x_0 + (\exp A(t - t_0) - Id) A^{-1}b \quad (2.32)$$

Thus, the behaviors of (2.27) and (2.29) are entirely driven by the “contraction” and “expansion”¹¹ of the exponential $\exp(At)$. Let us recall that:

$$\exp(At) = \sum_{i=0}^{\infty} \frac{(At)^i}{i!} = \sum_{i=0}^{n-1} \alpha_i(t) A^i \quad (2.33)$$

¹¹“Contractions” along the eigenvectors corresponding to the eigenvalues of A with negative real parts and “expansions” along the eigenvectors corresponding to the eigenvalues of A with positive real parts.

since, from the Cayley–Hamilton Theorem, A^n is a linear combination of the A^i , $0 \leq i \leq (n-1)$. There exist various ways to compute the exponential, some of which are recalled below:

1. Rewriting A using its Jordan canonical form:

$$A = PJP^{-1}, \quad J = \text{diag}(J(\lambda_i))$$

$$J(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_i \end{pmatrix}$$

thus $\exp(At) = P \exp(Jt)P^{-1}$ with:

$$\exp(Jt) = \text{diag}(\exp(J(\lambda_i)t))$$

$$\exp(J(\lambda_i)t) = \exp(\lambda_i t) \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^{(k-1)}}{(k-1)!} \\ 0 & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & \ddots & \ddots & t \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

2. Using the Dunford splitting: $A = N + D$, with N nilpotant ($N^n = 0$) and D diagonalisable (in $\mathbb{C}$), since $\exp(At) = \sum_{i=0}^{\infty} ((N + D)t)^i / i!$ and $(N + D)^i = \sum_{k=0}^i C_k^i N^{i-k} D^k$. The computation of the exponential is then simplified since $N^n = 0$: this trick is similar to the first one and is sometimes faster.
3. By using the method of the constituting matrices, the matrix $f(A)$ can be given by:

$$f(A) = \sum_{i=1}^r \sum_{j=0}^{n_i-1} f^{(j)}(\lambda_i) Z_{ij}$$

where r is the number of distinct eigenvalues of A denoted by λ_i and n_i their multiplicity. Thus, $f(A)$ is written as a linear combination of matrices Z_{ij} independent of the function f (which depends only on A); the coefficients of this combination depend, then, on the function f . Thus, one determines Z_{ij} using a simple testing function (e.g., $x \mapsto x^k$).

For (2.29), in the particular case of b constant:

$$\dot{x} = Ax + b \quad (2.34)$$

the equilibrium points satisfy $Ax + b = 0$. If A is regular, then there is only one equilibrium point given by $x_e = -A^{-1}b$. If A is singular, then two cases arise:

- There is an infinite number of equilibrium points if $\text{rank}(A, b) = \text{rank}(A)$, that is, $b \in \text{image}(A)$ or, if $\exists c \in \mathbb{R}^n : b = Ac$. They are then defined by: $x_e = c + u_0$, where $u_0 \in \ker(A)$ (u_0 is any eigenvector of A associated to the null eigenvalue).
- There is no equilibrium if $b \notin \text{image}(A)$.

2.4.2 Stability Conditions

THEOREM 7

Let A be an $n \times n$ -matrix with entries in $\mathbb{R}$, its spectrum being $\sigma(A) = \{\lambda_i \in \mathbb{C}, i = 1, \dots, r \leq n : \det(\lambda_i Id - A) = 0 \text{ and } \lambda_i \neq \lambda_j \text{ for } i \neq j\}$, and $v(\lambda_i)$ the smallest integer such that $\ker(A - \lambda_i I)^{v(\lambda_i)+1} = \ker(A - \lambda_i I)^{v(\lambda_i)}$. Let b be a constant vector satisfying $b \in \text{image}(A)$ and $x_e : Ax_e + b = 0$.

1. If $\exists \lambda_i \in \sigma(A) : \text{Re}(\lambda_i) > 0$, then $\lim_{t \rightarrow +\infty} \|\exp(At)\| = +\infty$ and x_e is unstable for (2.34).
2. If $\exists \lambda_i \in \sigma(A) : \text{Re}(\lambda_i) = 0$ and $v(\lambda_i) > 1$, then $\lim_{t \rightarrow +\infty} \|\exp(At)\| = +\infty$ and x_e is unstable for (2.34).
3. If $\text{Re}(\lambda_i) < 0, i = 1, \dots, (r-1)$ and $\text{Re}(\lambda_r) = 0$, with $v(\lambda_r) = 1$, then $\|\exp(At)\| < +\infty$ and x_e is stable but non-attractive for (2.34).
4. If $\forall \lambda_i \in \sigma(A) : \text{Re}(\lambda_i) < 0$, then $\lim_{t \rightarrow +\infty} \|\exp(At)\| = 0$ and x_e is exponentially stable (thus asymptotically stable) for (2.34).

REMARK 4

Since there is no possible confusion due to equilibriums of (2.34) all having the same stability property, one also speaks about the “stability of the system (2.34),” or “of the matrix A .”

Example 7

The equilibrium $x_e = 0$ of (2.27) is unstable for

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ (Case 1)}$$

and for

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ (Case 2)}$$

stable for

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ (Case 3)}$$

and exponentially stable for

$$A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \text{ (Case 4).}$$

Notice that, in the second and third case, A has two null eigenvalues. The conclusion is obtained according to $\nu(0)$ which is 2 in the second case and 1 in the third.

From this, one deduces the following necessary and sufficient condition for the origin to be asymptotically stable for an autonomous linear system.

COROLLARY 1

x_e is asymptotically stable for (2.34) $\Leftrightarrow \forall \lambda_i \in \sigma(A) : \operatorname{Re}(\lambda_i) < 0$. In this case, $x_e = -A^{-1}b$ and the stability is also exponential.

COROLLARY 2

If the characteristic polynomial of A is as follows $\pi_A(x) = x^n + \sum_{i=0}^{n-1} a_i x^i$, then a necessary condition of stability of x_e for (2.34) is that all a_i are positive.

Note that a necessary and sufficient condition of asymptotic stability is that $\pi_A(x)$ is Hurwitz, or that the a_i satisfies the Routh criterion [13].

Other necessary and sufficient conditions are available in the autonomous linear case: some are based on the Liapunov equation, others relate to only the particular shapes of matrix A : if $A = (a_{ij})$ with $a_{ij} \geq 0$ for any $i \neq j$, then A is asymptotically stable if and only if the principal minors of $-A$, that is, the n cascaded determinants of the matrices

$$(-a_{11}), \det \begin{pmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{pmatrix}, \dots, \det(-A)$$

are all positive (Kotelianskii criterion).

2.5 Behavior Studies: Local Results

This section deals with nonlinear autonomous systems of the form (2.17), for which the existence and uniqueness of solution to the associated CP is assumed to be $\mathcal{I}(x_0) = \mathcal{I} = \mathbb{R}$.

2.5.1 Structural Stability

It is a well-known fact that the real world cannot be reduced to a mathematical model: any modelling activity leads to some incorrect terms or models. This is why a natural question arises regarding for what kind of function $p(x)$ do the solution of (2.17) and that of:

$$\frac{dx}{dt} = g_p(x, p(x)), \quad x \in \mathcal{X} \quad (2.35)$$

look the same. To study this property, called *structural stability*, we need to introduce some distance over the set of $\mathcal{C}^1$ vector fields, which allows to characterize the closeness of two vector fields, together with the notion of *topological equivalence* which allows to compare the “resemblance” (likeness) of the solutions. This structural stability property is well known from people in control because it is a kind of robustness. The following definitions are taken from p. 91–140 of [1], p. 38–42 of [12], p. 305–318 of [15], and p. 93–101 of [22].

2.5.1.1 Whitney Distance

DEFINITION 23 Let g_1 and g_2 be two $\mathcal{C}^1(\mathcal{X})$ vector fields. The Whitney distance or $\mathcal{C}^1$ distance on $\mathcal{S} \subset \mathcal{X}$ between g_1 and g_2 , is defined by:

$$\rho_S^1(g_1; g_2) = \max \left\{ \sup_{x \in \mathcal{S}} \left\{ \left\| \frac{\partial^j g_{1i}(x)}{\partial x^j} - \frac{\partial^j g_{2i}(x)}{\partial x^j} \right\| : i = 1, \dots, n \right\} : j = 0, 1 \right\}.$$

For g , a $\mathcal{C}^1(\mathcal{X})$ vector field, one defines an ε -neighborhood of g in the $\mathcal{C}^1$ sense on the set $\mathcal{S} \subset \mathcal{X}$ as the set of all $\mathcal{C}^1(\mathcal{X})$ vector fields g' satisfying $\rho_S^1(g'; g) \leq \varepsilon$.

2.5.1.2 Equivalence

DEFINITION 24 h is a *conjugacy* with respect to $\mathcal{S} \subset \mathcal{X}$ between the solutions of (2.17) and (2.35) iff: $\forall x_0 \in \mathcal{S}$, $h[\Phi_{g_p}^t(x_0)] = \Phi_g^t[h(x_0)]$ holds for all time $t \in \mathbb{R}$ for which $\Phi_{g_p}^t(x_0)$ and $\Phi_g^t[h(x_0)]$ live in $\mathcal{S}$.

This relation just says that the function h maps the orbits of the perturbed system (2.35) onto the orbits of the unperturbed system (2.17).

DEFINITION 25 Systems (2.17) and (2.35) are *topologically* (resp. *differentially*, *linearly*) equivalent with respect to $\mathcal{S}$ iff there exists h continuous (resp. differentiable, linear) conjugacy with continuous (resp. differentiable, linear) inverse and which maps the solution of (2.17) onto that of (2.35) (within the set $\mathcal{S}$).

Since $[h \text{ linear}] \Rightarrow [h \text{ differentiable}] \Rightarrow [h \text{ continuous}]$, the topological equivalence is the weakest notion:

$$[\text{linear equivalence}] \Rightarrow [\text{differentiable equiv.}] \Rightarrow [\text{topological equiv.}]$$

2.5.1.3 Structural Stability

DEFINITION 26 System (2.17) is structurally stable with respect to $\mathcal{S}$ if there exist an ε -neighborhood of g in the $\mathcal{C}^1$ sense defined on $\mathcal{S} \subset \mathcal{X}$, such that, for all vector fields g_p from this neighborhood, system (2.35) associated to g_p and system (2.17) are topologically equivalent.

2.5.2 From Linear Model to Nonlinear Model

This section presents the connections between local behavior of autonomous nonlinear system described by (2.17) and the following linear ODE:

$$\frac{dx}{dt} = Ax, \quad x \in \mathcal{X} \quad (2.36)$$

under the assumption:

$$(H) \quad \begin{cases} \mathcal{X} = \mathbb{R}^n \\ \text{and } \forall x_0 \text{ initial condition there exist a unique} \\ \text{maximal solution defined on } \mathcal{I}(x_0) = \mathcal{I} = \mathbb{R}. \end{cases}$$

The techniques to study local behaviors are based on a fundamental result on structural stability (Section 2.5.2.1), which brings the local study of (2.17) to that of a system like (2.36). This local study is made around critical elements like equilibrium points or closed orbits.

2.5.2.1 Structural Stability Theorem and Consequences

When looking at (2.36), one has to partition the spectrum of A (denoted by $\sigma(A)$) into three parts:

$$\begin{cases} \sigma_s(A) = \{\lambda \in \sigma(A) : \operatorname{Re}(\lambda) < 0\} \\ \sigma_c(A) = \{\lambda \in \sigma(A) : \operatorname{Re}(\lambda) = 0\} \\ \sigma_u(A) = \{\lambda \in \sigma(A) : \operatorname{Re}(\lambda) > 0\} \end{cases}$$

where the indexes s, c, u mean, respectively, “stable,” “center,” and “unstable.” Then the corresponding generalized eigenvectors of A are used to obtain the following subspaces $E_s(A)$, $E_c(A)$, $E_u(A)$, whose dimensions are, respectively, n_s , n_c , n_u with:

$$\begin{aligned} n_s + n_c + n_u &= n \\ E_s(A) \oplus E_c(A) \oplus E_u(A) &= \mathbb{R}^n \end{aligned}$$

Note that a similar partition is possible when $\mathcal{X}$ is n -dimensional manifold, since $\mathcal{X}$ is locally the “same” as $\mathbb{R}^n$. When $\sigma_c(A) = \emptyset$, A is hyperbolic (see Definition 8). Similarly, (2.36) is asymptotically stable if and only if $\sigma(A) = \sigma_s(A)$, $\sigma_c(A) = \sigma_i(A) = \emptyset$; A is then said asymptotically stable or Hurwitz.

THEOREM 8 (First Liapunov method)

Let x_e be an equilibrium point of (2.17), to which the linearized model (2.36) is associated:

1. $\sigma_u(A) = \sigma_c(A) = \emptyset \Rightarrow x_e$ is asymptotically stable for (2.17)
2. $\sigma_u(A) \neq \emptyset \Rightarrow x_e$ is unstable for (2.17)

Thus, if the origin is asymptotically stable for the linearized model, then the corresponding equilibrium point is also asymptotically stable for the original nonlinear model.

COROLLARY 3

Under Assumption (H), and if A is hyperbolic, then (2.36) is structurally stable.

Then A is said to be structurally stable: $[A \text{ is hyperbolic}] \Leftrightarrow [A \text{ is structurally stable}]$. Note that Assumption (H) is of prime importance, as example (2.14) shows. A direct consequence is the following (p. 99 of [22]).

THEOREM 9

Assume that (H) holds, if (2.17) has an hyperbolic equilibrium point x_e , then there exists a neighborhood $\mathcal{V}(x_e)$ of x_e such that (2.17) is structurally stable within the set $\mathcal{V}(x_e)$.

2.5.2.2 Local Structure of Solution Within a Neighborhood of an Equilibrium Point

For a hyperbolic equilibrium point, the following results gives some insight about the local behavior of the solution of (2.17) around this point.

THEOREM 10 (Hartman–Grobman, 1964)

If the jacobian matrix $J_g(x_e) = A$ evaluated at the equilibrium point x_e does not have a purely imaginary or null eigenvalue ($\sigma_c(A) = \emptyset$), then there exists a homeomorphism h defined on $\mathcal{V}(x_e)$ a neighborhood of x_e , locally mapping the orbits of the linear flow of (2.36) onto those of the nonlinear flow Φ_g^t of (2.17). Moreover, h preserves the going-through direction on the orbits and can be selected in order to preserve the time parameterization.

From the neighborhood $\mathcal{V}(x_e)$ on which h is defined, one builds the *stable and unstable local manifolds*:

$$W_{\text{loc } s}(x_e) = \{x \in \mathcal{V}(x_e) : \lim_{t \rightarrow +\infty} \Phi_g^t(x) = x_e \text{ and } \Phi_g^t(x) \in \mathcal{V}(x_e), \forall t > 0\}$$

$$W_{\text{loc } u}(x_e) = \{x \in \mathcal{V}(x_e) : \lim_{t \rightarrow -\infty} \Phi_g^t(x) = x_e \text{ and } \Phi_g^t(x) \in \mathcal{V}(x_e), \forall t > 0\}$$

from which one defines stable and unstable manifolds (with respect to x_e):

$$W_s(x_e) = \cup_{t \geq 0} \Phi_g^t(W_{\text{loc } s}(x_e))$$

$$W_u(x_e) = \cup_{t \leq 0} \Phi_g^t(W_{\text{loc } u}(x_e))$$

These concepts of stable and unstable manifolds thus exhibit solutions of (2.17) which are respectively “contracting” and “expanding.” The manifolds $W_s(x_e)$, $W_u(x_e)$ are images by h of the corresponding subspaces on the linearized model: $W_s(x_e) = h[E_s(J_g(x_e))]$, $W_u(x_e) = h[E_u(J_g(x_e))]$.

THEOREM 11 (Stable manifold)

If (2.17) has a hyperbolic equilibrium point x_e , then there exists $W_s(x_e)$ and $W_u(x_e)$:

1. *Of dimension n_s and n_u the same as those of the spaces $E_s(J_g(x_e))$ and $E_u(J_g(x_e))$ of the linearized system (2.36) (with $A = J_g(x_e)$)*
2. *Tangents to $E_s(J_g(x_e))$ and $E_u(J_g(x_e))$ at x_e*
3. *Invariant by the flow Φ_g^t*

Moreover, $W_s(x_e)$ and $W_u(x_e)$ are manifolds as smooth as g (of the same class r as $g \in \mathcal{C}^r(\mathbb{R}^n)$).

In the critical case of nonhyperbolic points (or degenerated), the following result holds (see p. 127 of [12]).

THEOREM 12 (Center manifold) (Kalley, 1967)

Let g be a $C^r(\mathbb{R}^n)$ vector field, admitting a degenerated equilibrium point x_e . Let us denote as $A = J_g(x_e)$ its Jacobian matrix evaluated at this point. Then, there exist:

1. $W_s(x_e)$ and $W_u(x_e)$ invariant manifolds called, respectively, stable and unstable of class C^r , tangent to $E_s(J_g(x_e))$ and $E_u(J_g(x_e))$ at x_e
2. $W_c(x_e)$ a center manifold of class $C^{(r-1)}$ tangent to $E_c(J_g(x_e))$ at x_e

The manifolds $W_s(x_e)$, $W_u(x_e)$, and $W_c(x_e)$ are all invariant by the flow Φ_g^t and of the same dimension, as the corresponding subspaces of the linearized system (2.36) ($E_s(J_g(x_e))$, $E_u(J_g(x_e))$, and $E_c(J_g(x_e))$). The stable and unstable manifolds ($W_s(x_e)$ and $W_u(x_e)$, respectively) are unique, whereas $W_c(x_e)$ is not necessarily so.

However, it is difficult to obtain these manifolds, even numerically: often, the only recourse for the determination of a center manifold is to use a Taylor expansion of $W_c(x_e)$ in the neighborhood of the degenerated point x_e : this method has been known for a long time since A.M. Liapunov used it in 1892 to study the “critical case” [19].

For the sake of simplicity, one carries out a change of coordinates on the initial system (2.17) to come back to the case with the equilibrium point being at the origin. In the most interesting case in practice, $W_u(0)$ is empty. The center manifold theorem ensures that the initial system (2.17) is topologically equivalent to:

$$\begin{cases} \frac{dx_c}{dt} = A_c x_c + g_1(x) \\ \frac{dx_s}{dt} = A_s x_s + g_2(x) \end{cases}$$

with A_c of dimension n_c corresponding to that of $E_c(J_g(0))$ and which thus has all its eigenvalues with null real part. A_s is of dimension n_s corresponding to that of $E_s(J_g(0))$, therefore asymptotically stable. One can express $W_c(0)$ as an hypersurface:

$$W_c(0) = \{(x_c, x_s) \in \mathbb{R}^{n_c} \times \mathbb{R}^{n_s} : x_s = k(x_c)\}$$

Moreover, one knows that when $W_c(0)$ contains 0 (thus $k(0) = 0$) and, at this point, is tangential to $E_c(J_g(0))$ (thus $J_k(0) = 0$). One gets:

$$x_s = k(x_c) \implies \frac{dx_s}{dt} = J_k(x_c) \frac{dx_c}{dt}$$

thus:

$$A_s x_s + g_2(x_c, k(x_c)) = J_k(x_c)(A_c x_c + g_1(x_c, k(x_c))) \quad (2.37)$$

$$k(0) = 0, \quad J_k(0) = 0 \quad (2.38)$$

One studies the projection of the vector field of $x_s = k(x_c)$ onto $E_c(J_g(0))$:

$$\frac{dx_c}{dt} = A_c x_c + g_1(x_c, k(x_c)) \quad (2.39)$$

taking into account (2.37) and (2.38). This leads to the following theorem (see p. 131 of [12]).

THEOREM 13 (Henry and Carr, 1981)

If:

1. $W_u(0)$ is empty
2. the equilibrium $x_{ec} = 0$ of (2.39) is locally asymptotically stable (resp. unstable),

then the equilibrium x_e of (2.17) is asymptotically stable (resp. unstable).

The computation of (2.39) being generally impossible, the following theorem [12] makes possible the study of local stability of the equilibrium $x_{ec} = 0$ by using the approximate of k .

THEOREM 14 (Henry and Carr, 1981)

If there exist $\psi : \mathbb{R}^{n_c} \rightarrow \mathbb{R}^{n_s}$ with $\psi(0) = 0$ and $J_\psi(0) = 0$, such that, when $x \rightarrow 0$:

$$J_\psi(x_c)[A_c x_c + g_1(x_c, \psi(x_c))] - A_s \psi(x_c) - g_2(x_c, \psi(x_c)) = o(x^r), \quad r > 1 \quad (2.40)$$

then $h(x_c) = \psi(x_c) + o(x^r)$, when $x \rightarrow 0$.

This technique allows, in most of the cases, to arrive at a conclusion about the asymptotic stability of a degenerated equilibrium.

Example 8

Let us consider the following ODE $(x, y) \in \mathbb{R}^2$:

$$\frac{dx}{dt} = -x^2 + xy \quad (2.41)$$

$$\frac{dy}{dt} = -y + x^2$$

One gets:

$$J_g(x, y) = \begin{pmatrix} -2x + y & x \\ 2x & -1 \end{pmatrix}$$

and the system has two equilibrium points:

$$z_{e1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ degenerated, } J_g(z_{e1}) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

$$z_{e2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ unstable, } J_g(z_{e2}) = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

For the origin, the eigenvalues of the associated jacobian matrix are 0 and -1 (one gets $A_c = 0, A_s = -1$). One looks for the center manifold associated to this equilibrium point by his third-order development: $k(x) = ax^2 + bx^3 + o(x^3)$, since $k(0) = J_k(0) = 0$. This development must satisfy (2.40): therefore,

$$[2ax + 3bx^2 + o(x^2)][-x^2 + (ax^3 + bx^4 + o(x^4))] = [(1 - a)x^2 - bx^3 + o(x^3)]$$

and, by equalizing the terms of the same degree, one obtains $a = 1, b = 2$, this is: $k(x) = x^2 + 2x^3 + o(x^3)$. Thus (2.39) becomes $\dot{x} = -x^2 + x^3 + o(x^3)$ and Theorem 13 makes it possible to conclude with unstability the origin. Notice that the same result can be obtained more intuitively and without too much calculation, by noting that the second dynamics (in y) of (2.41) converges faster (exponentially) than the first one (in x): one can thus consider that after a transient, $(dy/dt) = 0 = -y + x^2$, (i.e., $y = x^2$): one finds the center manifold $k(x) = x^2 + o(x^2)$. This is justified using the singular perturbation theory theorem.

Example 9

Let us consider the following ODE:

$$\frac{dx}{dt} = xy$$

$$\frac{dy}{dt} = -y - x^2$$

with $(x, y) \in \mathbb{R}^2$. The origin is the only equilibrium point. The eigenvalues of the associated jacobian matrix are 0 and -1 . A third-order development of $k(x)$ is $-x^2 + o(x^3)$. Theorem 13 leads to the conclusion that the origin is asymptotically stable (but not exponentially stable).

REMARK 5

There exists a quick way to deal with these two examples: in the neighborhood of the origin, y converges exponentially, thus “infinitely faster” than x would do it. One deduces from this that dy/dt cancels “infinitely” faster than dx/dt ; this is, for Example 8, $y = x^2$, which deferred in $dx/dt = -x^2 + y$ gives the approximate equation $dx/dt = -x^2 + x^3$. In the same way, Example 9, when $t \rightarrow \infty$, leads to $y \rightarrow -x^2$, therefore $dx/dt = -x^3$.

2.5.2.3 Local Structure of the Solutions in the Neighborhood of a Closed Orbit

To study the local structure of a closed orbit, one introduces the concept of the Poincaré¹² section which makes it possible to define an application known as the Poincaré map. The map then brings back the local study of a closed orbit, for a continuous dynamic system, to the local study of a fixed point for a discrete dynamic system.¹³ This theoretical tool can be used in practice only using numerical algorithms. This process outlined here is based on pp. 243, 281–285 of [15] and p. 23–27 of [12].

Assume that the dynamic system (2.17) has a closed orbit γ (see Definition 10) and let x_γ be a point of this orbit.

DEFINITION 27 S_γ is a local Poincaré section at x_γ of (2.17) if:

1. S_γ is an open of a submanifold $\mathcal{V}$ of $\mathcal{X}$ having dimension $(n - 1)$ and containing x_γ
2. $T\mathcal{V}(x_\gamma)$ tangent space to $\mathcal{V}$ at x_γ and $g(x_\gamma) \in \mathbb{R}^n$ are in direct sum:
 $\mathbb{R}^n = T\mathcal{V}(x_\gamma) \oplus g(x_\gamma)\mathbb{R}$

This last condition expresses the transversality (nontangent) of S_γ and the vector field $g(x)$ of (2.17) at x_γ . Let S_γ be a local Poincaré section (Figure 2.12) at x_γ of (2.17): since $x_\gamma \in \gamma$, one gets $\Phi_g^T(x_\gamma) = x_\gamma$, denoting by T_γ the period of γ . If one considers x_0 a point sufficiently close to x_γ , there exist a time $T(x_0)$ close to T_γ after which $\Phi_g^{T(x_0)}(x_0) \in S_\gamma$. Let us consider $\mathcal{V}(x_\gamma)$ a neighborhood of x_γ and let us build the map known as Poincaré map or the *first return map*:

$$P : \mathcal{V}(x_\gamma) \cap S_\gamma \longrightarrow S_\gamma$$

$$x_0 \mapsto P(x_0) = \Phi_g^{T(x_0)}(x_0)$$

¹²Henri Poincaré (1854–1912), French mathematician and physicist. Entered the Polytechnic School in 1873, and became an engineer from the “cops des Mines” in 1877. Then he taught at the Faculty of Science of Caen and then at the Sorbonne in 1881.

¹³The just seen results concerning the equilibrium points of an ODE (2.17) can be transposed to the fixed points for an discrete equation of recurrence of $x_{k+1} = g(x_k)$.

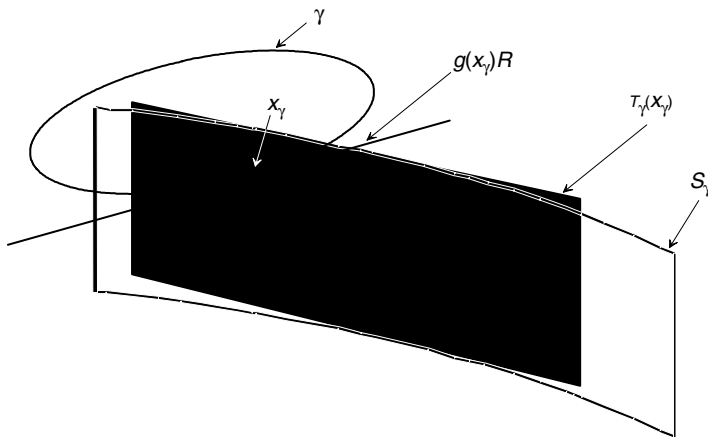


FIGURE 2.12
Poincaré section.

This construction is justified since under some assumptions [e.g., g of class $\mathcal{C}^1(\mathbb{R}^n)$] there exists $\mathcal{V}(x_\gamma)$ and a unique application $T : \mathcal{V}(x_\gamma) \rightarrow \mathbb{R}, x_0 \mapsto T(x_0)$, such that $\forall x_0 \in \mathcal{V}(x_\gamma) : \Phi_g^{T(x_0)}(x_0) \in S_\gamma$ and $T(x_\gamma) = T_\gamma$. Note that P depends neither on x_γ nor on S_γ . P generates a discrete dynamic system and has a fixed point x_γ ($P(x_\gamma) = x_\gamma$). Thus, this application brings back the study of the solutions in the neighborhood of a closed orbit of a continuous dynamic system defined on a manifold $\mathcal{X}$ of dimension n to the study of the solutions in the neighborhood of a fixed point of a discrete dynamic system defined on a manifold of dimension $n - 1$: $x_{k+1} = P(x_k) = P^k(x_0)$. The local behavior of the solutions of the discrete dynamic system in the neighborhood of the fixed point fixes x_γ makes it possible to deduce the behavior of the solutions for the continuous dynamic system (2.17) in the neighborhood of γ .

The study of a the local behavior of a discrete dynamic system in a neighborhood of a fixed point is very similar to the study of a continuous dynamic system in a neighborhood of an equilibrium point. In particular, to study the discrete system $x_{k+1} = Ax_k$, one partitions $\sigma(A)$ into $\sigma_s(A) = \{\lambda \in \sigma(A) : |\lambda| < 1\}$, $\sigma_c(A) = \{\lambda \in \sigma(A) : |\lambda| = 1\}$, $\sigma_u(A) = \{\lambda \in \sigma(A) : |\lambda| > 1\}$. One then obtains similar results to those previously developed, allowing to deduce from it the local structure of the flow in the neighborhood of a closed orbit γ .

However, the application P can be obtained explicitly only if the solutions of (2.17) can be explicitly computed: this limits the practical interest of the application P which very often must be evaluated numerically.

Example 10

Using polar coordinates ($x_1 = r \cos(\theta)$, $x_2 = r \sin(\theta)$), ODE (2.13) becomes:

$$\begin{aligned}\frac{dr}{dt} &= r(1 - r^2) \\ \dot{\theta} &= 1\end{aligned}$$

Solutions are $r(t) = r_0 / \sqrt{r_0^2(1 - e^{-2t}) + e^{-2t}}$, $\theta(t) = \theta_0 + t$, and straightforward study shows that any solution starting in the plane except the origin converge to a periodic solution $x_1(t) = \cos(t + \phi)$, $x_2(t) = \sin(t + \phi)$. Instead of this direct analysis let us use the Poincaré section $x_2 = 0$, in the neighborhood of $(1, 0)$. One brings back the study to $x_1^{k+1} = x_1^k / \sqrt{(x_1^k)^2(1 - e^{-4\pi}) + e^{-4\pi}}$ whose linearized model is $y_1^{k+1} = e^{-4\pi} y_1^k$. Thus one concludes that the periodic orbit is locally asymptotically stable.

2.6 Bifurcation and Chaos

Nonlinear models can have radical changes of behavior when a parameter evolves: this is a bifurcation phenomenon. For example, the displacement of a mass m attached to a spring of stiffness k and a frame excited by force $\alpha \dot{x}$ is modeled by $m\ddot{x} + \mu\dot{x} + kx = 0$, $\mu = \delta - \alpha$, with δ the friction coefficient. The modes, for μ small, are $\lambda = (-\mu \pm i\sqrt{4mk - \mu^2})/2m$. Obviously, if μ is positive (resp. negative), then bottom equilibrium is unstable (resp. stable), whereas, for $\mu_0 = 0$, an oscillatory mode appears. Clearly, μ_0 is a bifurcation value.

As examples of simple discrete equations,¹⁴ one can check that an infinite number of such bifurcations can lead to an unforeseeable behavior due to their high sensitivity to the initial conditions: it is a phenomenon of chaos. This same type of phenomenon appears for autonomous nonlinear ODEs, but only for the dynamic of order equal to or higher than three.

2.6.1 Parameter Dependence

Assume that k parameters, gathered in a vector $\mu \in \mathbb{R}^k$, appear in the ODE:

$$\frac{dx}{dt} = g(x, \mu), \quad x \in \mathbb{R}^n, \mu \in \mathbb{R}^k \quad (2.42)$$

¹⁴For example, first-order discrete equation $x_{n+1} = \mu x_n(1 - x_n)$ [3, 12].

When the vector field g is of class $\mathcal{C}^1$ on open $\mathbb{R}^{n+k}$, in addition to “the existence, the continuity and the uniqueness of the solutions,” one can note (see Theorem 4) that for a given triplet (t_0, x_0, μ_0) , there are two open neighborhoods of x_0 and μ_0 noted, respectively, $\mathcal{V}(x_0)$ and $\mathcal{V}(\mu_0)$, such that for any pair (x_{01}, μ_{01}) in $\mathcal{V}(x_0) \times \mathcal{V}(\mu_0)$, the CP: $[dx/dt = g(x, \mu_{01}), x(t_0) = x_{01}]$ has one and only one maximum solution $\phi(t; t_0, x_{01}, \mu_{01})$ of $\mathcal{C}^1$ class with respect to t , x_{01} , and μ_{01} . In addition, under the terms of Theorem 6 (since solutions are unique), the equilibrium points of (2.42) are given by the solutions of:

$$g(x, \mu) = 0, \quad x \in \mathbb{R}^n, \mu \in \mathbb{R}^k$$

Thus, when the vector parameter μ varies, the implicit function theorem shows that these equilibriums $x_e(\mu)$ are related to μ with the function of the same class as g , provided that there is a equilibrium point (x_e, μ_e) and that the jacobian matrix evaluated at this point $J_g(x_e, \mu_e) = (\partial g / \partial x)(x_e, \mu_e)$ is nonsingular: $\det((\partial g / \partial x)(x_e, \mu_e)) \neq 0$. Under these conditions, there is an open neighborhood of μ_e , noted $\mathcal{V}(\mu_e)$ and an application $h : \mathcal{V}(\mu_e) \subset \mathbb{R}^k \rightarrow \mathbb{R}^n, \mu \mapsto h(\mu)$ of the same class as g , such that:

$$g(h(\mu), \mu) = 0, \quad \forall \mu \in \mathcal{V}(\mu_e)$$

DEFINITION 28 The graph of the function h constitutes *branches of equilibriums*.

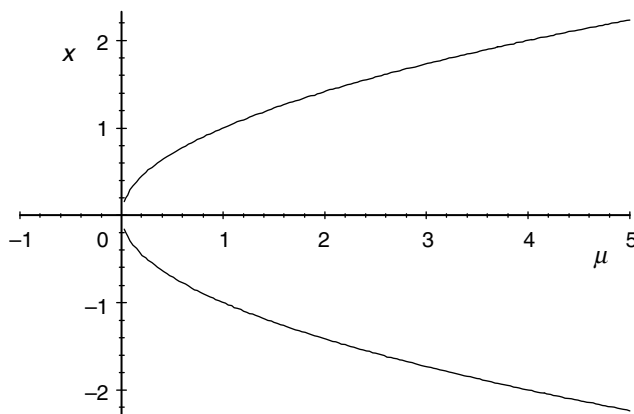
Example 11

Let us consider the ODE: $dx/dt = \mu^3 - x^3$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$. The branch of equilibriums is the line $x = \mu$. Note that, for any equilibrium $(x, \mu) \neq (0, 0)$, the jacobian matrix of g is nonsingular: in this example, the mapping h is the identity or its opposite according to the sign of $x_e \mu_e$.

Example 12

For the ODE $dx/dt = \mu - x^2$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$, the branches of equilibriums correspond to the graph of the parabola $x = \pm\sqrt{\mu}$, $\mu \geq 0$.

On the one hand, when $(\partial g / \partial x)(x_e, \mu_e)$ is nonsingular, the equilibrium points (in a neighborhood of (x_e, μ_e)) are hyperbolic: Theorem 11 can be used to study the local structure of the solutions in the neighborhood of these points (structurally stable system) (Figure 2.13). On the other hand, when $(\partial g / \partial x)(x_e, \mu_e)$ is singular, one is in the presence of a degenerated point (nonhyperbolic), which results in the possible presence of a change of behavior (junction). Note then that this point can be the junction of several branches of equilibriums (see Example 11). The condition “ $(\partial g / \partial x)(x_e, \mu_e)$ singular” induces a local bifurcation. A general definition of the concept of bifurcation is as follows.

**FIGURE 2.13**

Branch of equilibrium.

DEFINITION 29 A *bifurcation value* is a value of the vector parameter μ as in Equation (2.42), for which (2.42) is not structurally stable. In general, one distinguishes two kinds of bifurcations:

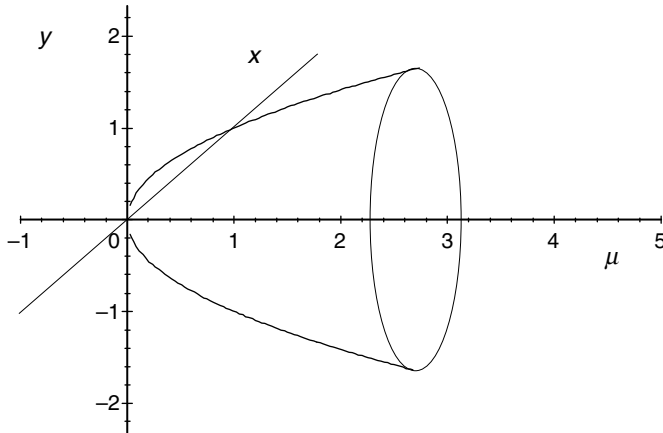
1. *Local bifurcations*: the qualitative changes of the phase portrait appear in the neighborhood of critical elements
2. *Global bifurcations*: the changes take place on a subspace of the state space, for example, when there is a creation of attractor strange, or when a homoclinic orbit is transformed into periodic orbit or into an equilibrium point

Example 13

$\mu_0 = 0$ is a bifurcation value for the ODE of Example 12, but not for that of Example 11 because the equilibrium $x = \mu$ is always asymptotically stable whatever be the value of the parameter μ .

DEFINITION 30 The graph, in the space (x, μ) , of the evolution of the invariants sets (equilibrium points, orbits closed, etc.) with respect to the parameter μ is a bifurcation diagram.

Here, the term “evolution” has to be understood in a qualitative sense, that is, it can be a question of creation or qualitative change (e.g., stable $\rightarrow$ unstable). Thereafter, the following convention will be adopted: the stable elements will be represented with straight lines and the unstable ones with discontinuous lines.

**FIGURE 2.14**

Hopf bifurcation of (2.3).

Example 14

Let us consider again the Van der Pol model (see Equation (2.3)). The origin is an equilibrium point and the jacobian matrix at this point is:

$$J_g(0) = \begin{pmatrix} 0 & 1 \\ -1 & 2\mu \end{pmatrix}$$

Thus, for μ near zero, the eigenvalues are $\mu \pm i\sqrt{1 - \mu^2}$, which means that $\mu_0 = 0$ is a bifurcation value for which the origin, while remaining an equilibrium point, qualitatively changes from “asymptotically stable” ($\mu < 0$) to “unstable” ($\mu > 0$). We will see hereafter that it is about a Hopf bifurcation which, when μ becomes positive, gives rise to an asymptotically stable limit cycle surrounding the origin. The bifurcation diagram is given in Figure 2.14.

2.6.2 Local Bifurcation Locale with Codimension 1

It is difficult to make an exhaustive classification of the phenomena of local or global bifurcation. A complex study appears with the increase of:

- The effective dimension of (2.42): n
- The number of parameters in (2.42): k

However, a great number of phenomena can be studied using “elementary bifurcations” that one often encounters.

In particular, for the equilibrium points, when the parameters vary, the eigenvalues of the jacobian matrix can cross the imaginary axis: this leads to bifurcation. Among these parameters, a minimal number can be used to reproduce this type of bifurcation: it is the codimension of the bifurcation.

For a bifurcation of codimension 1, $J_g(0)$ is similar to either

$$\begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix}$$

or

$$\begin{pmatrix} \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} & 0 \\ 0 & Y \end{pmatrix}$$

with X and Y matrices of respective size $(n-1) \times (n-1)$ and $(n-2) \times (n-2)$. Generically, any bifurcation (local in the neighborhood of an equilibrium) of codimension 1 can be reduced to one of the following bifurcation.¹⁵

2.6.2.1 Subcritical or Saddle–Node

This bifurcation is modeled by:

$$\frac{dx}{dt} = \mu - x^2, \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R} \quad (2.43)$$

The equilibrium points are $x_{e1} = -\sqrt{\mu}$ and $x_{e2} = \sqrt{\mu}$. For $\mu \geq 0$, their respective jacobians are $2\sqrt{\mu}$ and $-2\sqrt{\mu}$; $\mu_0 = 0$ is a bifurcation value. There is creation of two equilibrium points: “ x_{e1} does not exist” ($\mu < 0$) $\rightarrow$ “exists and is unstable” ($\mu > 0$) and “ x_{e2} does not exist” ($\mu < 0$) $\rightarrow$ “exists and is asymptotically stable” ($\mu > 0$). When $\mu = \mu_0 = 0$, (2.43) becomes $dx/dt = -x^2$, which admits as solutions $x(t) = x_0/(1 + (t - t_0)x_0)$, which shows that, when x_0 is positive, $x(t)$ converges towards zero and that, in the contrary, there is a finite time ($t_0 - (1/x_0)$) for which there is “explosion” ($x(t) = \infty$). The bifurcation diagram is given in Figure 2.15.

Note that the equilibrium points are not true “saddle points” and “nodes,” since it would be necessary for the state space to be of dimension 2. For that it is enough to associate with (2.43) the equation $dy/dt = -y$, $y \in \mathbb{R}$, which gives the bifurcation diagram of Figure 2.16. This bifurcation takes its full name “saddle–node”.

¹⁵For that, one will be able to use the center manifold theorem (Theorem 12): (see [12] for more details).

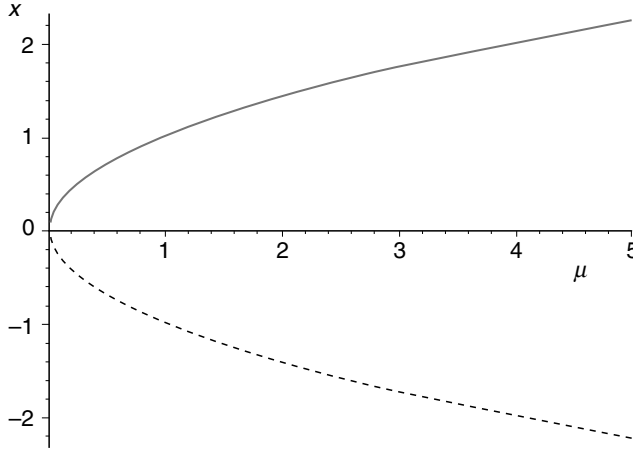


FIGURE 2.15
Saddle-node bifurcation of (2.43).

2.6.2.2 Transcritical Bifurcation

This bifurcation is modeled by:

$$\frac{dx}{dt} = \mu x - x^2, \quad x \in \mathbb{R}, \mu \in \mathbb{R} \quad (2.44)$$

The equilibrium points are $x_{e1} = 0$ and $x_{e2} = \mu$. Their respective jacobians are μ and $-\mu$; $\mu_0 = 0$ is a bifurcation value. There is exchange of stability

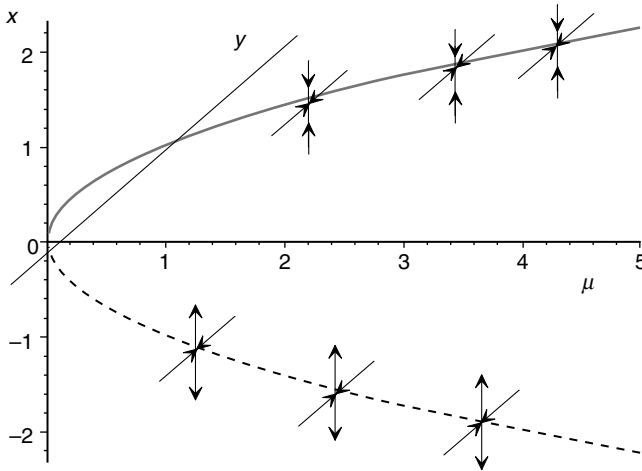
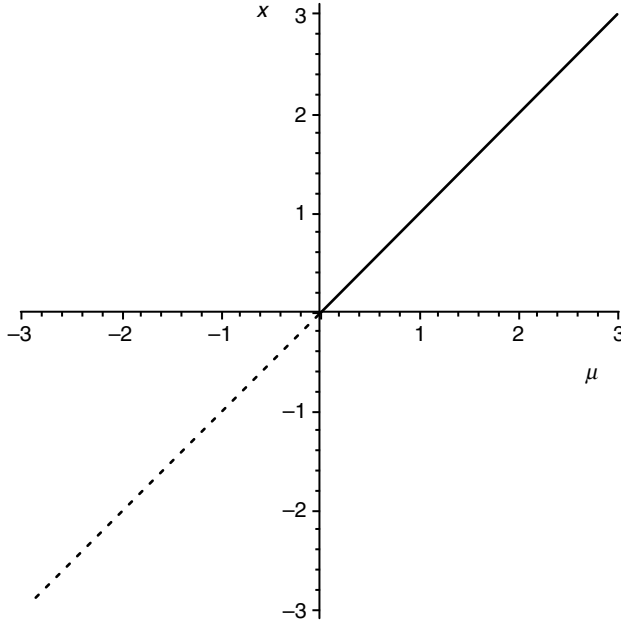


FIGURE 2.16
Saddle-node bifurcation with $\dot{y} = -y$.

**FIGURE 2.17**

Transcritical bifurcation of (2.44).

between the two equilibrium points: “ x_{e1} asymptotically stable” ($\mu < 0$) → “unstable” ($\mu > 0$) and “ x_{e2} unstable” ($\mu < 0$) → “asymptotically stable” ($\mu > 0$). When $\mu = \mu_0 = 0$, (2.44) becomes $dx/dt = -x^2$ (see earlier for the conclusions). The bifurcation diagram is given in Figure 2.17.

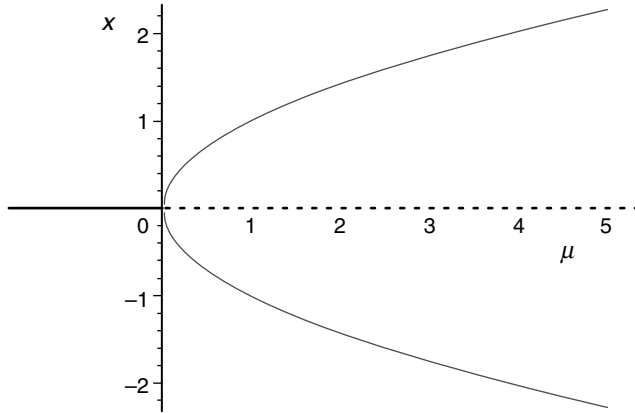
2.6.2.3 Supercritical

One distinguishes “*fork bifurcation*” and “*Hopf bifurcation*”.

The fork bifurcation is modeled by:

$$\frac{dx}{dt} = \mu x - x^3, \quad x \in \mathbb{R}, \mu \in \mathbb{R} \quad (2.45)$$

A quick study shows that $\mu_0 = 0$ is a bifurcation value, for which there is creation of two asymptotically stable equilibrium points and loss of stability for the origin: “ $x_{e1} = 0$ asymptotically stable” ($\mu < 0$) → “unstable” ($\mu > 0$) “ $x_{e2} = -\sqrt{\mu}$ does not exist” ($\mu < 0$) → “exists and is asymptotically stable” ($\mu > 0$) and “ $x_{e3} = \sqrt{\mu}$ does not exist” ($\mu < 0$) → “exists and is asymptotically stable” ($\mu > 0$). When $\mu_0 = 0$, (2.45) becomes $dx/dt = -x^3$, which admits as solution $x(t) = x_0 / \sqrt{1 + 2(t - t_0)x_0^2}$, which shows

**FIGURE 2.18**

Fork bifurcation of (2.45).

that $x(t)$ converges toward zero (the origin is asymptotically stable, not exponentially). The bifurcation diagram is given in Figure 2.18.

The Hopf bifurcation corresponds to the presence of two combined complex eigenvalues; it is modeled by:

$$\begin{aligned}\frac{dx}{dt} &= -\omega y + x(\mu - (x^2 + y^2)), & x \in \mathbb{R}, \mu \in \mathbb{R} \\ \frac{dy}{dt} &= +\omega x + y(\mu - (x^2 + y^2)), & y \in \mathbb{R}, \omega = \text{cste}\end{aligned}$$

This equation, in polar coordinates, becomes $dr/dt = r(\mu - r^2)$, $d\theta/dt = \omega$. These two equations are decoupled, the first corresponds to a fork bifurcation (valid only for r positive). Thus, one deduces from it that $\mu_0 = 0$ is a bifurcation value and that there is creation of an asymptotically stable closed orbit and loss of stability for the origin when μ becomes positive: origin: “asymptotically stable” ($\mu < 0$) $\rightarrow$ “unstable” ($\mu > 0$), orbit ($r = \sqrt{\mu}$): “does not exist” ($\mu < 0$) $\rightarrow$ “exists and is asymptotically stable” ($\mu > 0$). This leads to the Hopf bifurcation diagram given in Figure 2.14.

The mere presence of a parameter in an ODE does not mean the systematic existence of a bifurcation. Indeed, the ODE: $dx/dt = \mu - x^3$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ has only one equilibrium which is asymptotically stable for any value of μ : there is no bifurcation.

2.6.3 Chaos

A chaotic phenomenon (seemingly random behavior) can be obtained starting from several bifurcation phenomena: period doubling [3, 12, 20],

bifurcation on the torus (infinity of Hopf bifurcation), intermittency (periodic phenomena alternating with aperiodic phenomena), etc. The presence of a strange attractor is an indicator of chaos: indeed, this implies a great sensitivity of the solutions to the initial conditions (two solutions starting from close initial conditions give rise to trajectories of different natures or different forms). Also, a chaotic phenomenon can be detected by highlighting either an invariant set of “non-integer” size (strange attractor), or a sensitivity to the initial conditions (in particular using the Liapunov exponents). In what follows, one will consider only autonomous nonlinear ODEs of the type (2.17).

DEFINITION 31 A set $\mathcal{A}$ is strange attractor if $\mathcal{A}$ is an attractive invariant set by the flow Φ_g^t and if any trajectory initialized in $\mathcal{A}$ is dense in $\mathcal{A}$.

Example 15

Let us consider the Rössler model:

$$\begin{cases} \dot{x} = -(y + z) \\ \dot{y} = x + ay \\ \dot{z} = b - cz + xz \end{cases} \quad (2.46)$$

for $a = b = 0.2, c = 5.8$, one can get the Rössler attractor plotted in Figure 2.19.

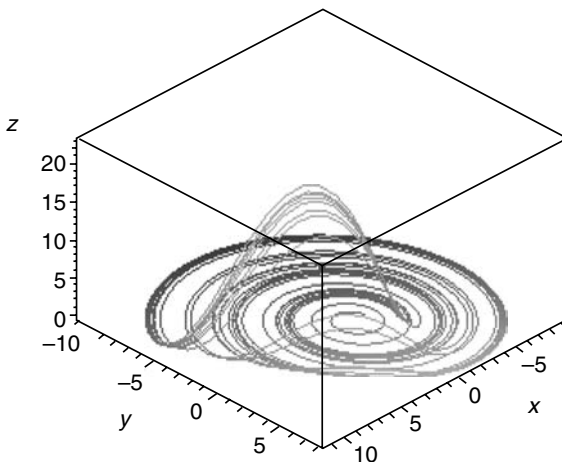


FIGURE 2.19
Rössler attractor.

From a practical point of view, it is very difficult to be able to show that a set $\mathcal{A}$ is a strange attractor, in particular to show that any trajectory initialized in $\mathcal{A}$ is dense in $\mathcal{A}$.

Also, it is natural to turn to numerical methods to make it possible to compute the dimension of an attractor which is a convincing indicator of sound “strangeness.” Consider a cube C containing attractor $\mathcal{A}$, whose dimensions need to be determined. Denoting by $n(\varepsilon)$ the number of cubes of edge ε necessary to cover all the points constituting the attractor $\mathcal{A}$, one defines *fractal dimension* (or *capacity*) by:

$$d_f(\mathcal{A}) = \lim_{\varepsilon \rightarrow 0} \frac{\ln(n(\varepsilon))}{\ln(1/\varepsilon)} \quad (2.47)$$

There are many other concepts of dimension: Hausdorff dimension (see p. 285 of [12]), entropy with respect to a measure (p. 286 of [12]), the information dimension (p. 345 of [26], p. 735 of [20]), the dimension of correlation (p. 345 of [26]), the Liapunov dimension (p. 739 of [20]), etc.¹⁶ However, in practice, one uses the fractal dimension which, from a numerical point of view, is obtained more easily. Indeed, by taking into account the precision of resolution of the ODE (see pp. 722–726 of [20]), it is enough to plot the curve $\ln(n(\varepsilon)) = f(\ln(1/\varepsilon))$ to obtain $d_f(\mathcal{A})$. For the Rösler attractor, one obtains $d_f(\mathcal{A}) = 2.015 \pm 0.005$.

Note that the strange attractor very often results from a process of “*feuilleter*”: a set is contracted in certain directions, is dilated in others, and is folded up on itself so that it is invariant (see the construction of the Smale horseshoe pp. 102–116 and 230–235 of [12] and pp. 328–334 of [26]). Thus to detect a strange attractor (thus a chaotic phenomenon), one can use the *Liapunov exponents* to measure the contractions (if the exponent is negative) and expansions (if the exponents is positive). Note that this characteristic results in a sensitivity to the initial conditions. Consider a ball of ray ε centered at a point x_0 ; then, the evolution of the axes of the reference frame $(\{e_i\}_{i=1,\dots,n})$ linked to this point is given by: $\Phi_g^t(x_0 + \varepsilon e_i)$, $i = 1, \dots, n$, allowing to define the i th Liapunov exponent by:

$$L_i = \lim_{t \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \frac{1}{t} \ln \left(\left| \frac{\Phi_g^t(x_0 + \varepsilon e_i)}{\varepsilon} \right| \right) \quad (2.48)$$

¹⁶All these dimensions can be defined starting from a parameterized family of dimension (known as Rényi dimension) defined by

$$d_q(\mathcal{A}) = \frac{1}{1-q} \lim_{\varepsilon \rightarrow 0} \frac{\ln \left(\sum_{i=1}^{n(\varepsilon)} p_i^q \right)}{\ln(1/\varepsilon)}, \quad q \geq 0$$

where p_i is the probability for a point of the attractor to be in the i th box. $n(\varepsilon)$ such boxes are needed to cover the whole attractor. Thus, if N is the number of points of the plotted attractor (obtained by simulation) and N_i is the number of points in the i th box, one gets $p_i = N_i/N$.

Consider a particular trajectory $\Phi_g^t(x_0)$ and denote by $\mu_i(t)$ the eigenvalues of the monodromy matrix¹⁷ of associated linearized model $\dot{x} = g(x)$. Then, around this trajectory (i.e., $\dot{z} = J_g(\Phi_g^t(x_0))z = A(t)z$), the Liapunov exponents are given by:

$$L_i = \lim_{t \rightarrow \infty} \frac{1}{t} \ln(|\mu_i(t)|)$$

In particular, if x_0 is an equilibrium point, then $\dot{z} = J_g(x_0)z = Az$. The monodromy matrix is $\Phi(t) = \exp(At)$; thus by noting λ_i the eigenvalues of A (all presumed real), one obtains $L_i = \lim_{t \rightarrow \infty} (1/t) \ln(\exp(\lambda_i t)) = \lambda_i$. In general, these Liapunov exponents are reordered as $L_1 \geq L_2 \geq \dots \geq L_n$. In this case, for a dissipative system $\sum_{i=1}^n L_i < 0$ and a necessary condition for the appearance of chaos is $L_1 > 0$. For an ODE of dimension 3, a necessary and sufficient condition for the existence of a strange attractor is: $L_1 < 0$, $L_2 = 0$, $L_3 > 0$.

References

1. V.I. Arnold, *Chapitres Supplémentaires À la Théorie Des Equations Différentielles Ordinaires*, MIR, Moscow, 1980.
2. V.I. Arnold, *Equations Différentielles Ordinaires*, MIR, Moscow, 1988, 4th ed., Russian translation.
3. P. Berge, Y. Pomeau, and CH. Vidal, *L'Ordre Dans Le Chaos (Vers Une Approche Déterministe de la Turbulence)*, 1984.
4. N.P. Bhatia and G.P. Szegő, *Stability Theory of Dynamical Systems*, Springer-Verlag, Berlin, 1970.
5. H.D. Chiang, M.W. Hirsch, and F.F. Wu, Stability regions of nonlinear autonomous dynamical systems, *IEEE Trans. Autom. Control*, 33 (1), 16–27, 1988.
6. H.D. Chiang and J.S. Thorp, Stability regions of nonlinear dynamical systems: a constructive methodology, *IEEE Trans. Autom. Control*, 34 (12), 1229–1241, 1989.
7. E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, 1955.
8. A.F. Filippov, *Differential Equations with Discontinuous Righthand Sides*, Kluwer Academic Publishers, 1988.
9. R. Genesio and A. Vicino, New techniques for constructing asymptotic stability regions for nonlinear systems, *IEEE Trans. Circuits Syst.*, CAS-31 (6), 574–581, 1984.

¹⁷This matrix is periodic in the case of a periodic trajectory.

10. R. Genesio, M. Tartaglia, and A. Vicino, On estimation of asymptotic stability regions: state of art and new proposals, *IEEE Trans. Autom. Control*, AC-30 (8), 747–755, 1985.
11. Lj.T. Grujić, A.A. Martynyuk, and M. Ribbens-Pavella, *Large Scale Systems Stability under Structural and Singular Perturbations*, LNCIS, Springer-Verlag, 1987.
12. J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, 1983.
13. W. Hahn, *Stability of Motion*, Springer-Verlag, N.Y., 1967.
14. J. Hale and H. Koçak, *Dynamics and Bifurcations*, vol. 3 of *Text in Applied Mathematics*, Springer-Verlag, N.Y., 1991.
15. M.W. Hirsh and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, 1974.
16. A. Isidori, *Nonlinear Control Systems*, 3rd ed., vol. 1, Springer, 1989.
17. H.K. Khalil, *Nonlinear Systems*, Prentice-Hall, 1996.
18. V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities*, vol. 1, Academic Press, New York, 1969.
19. A.M. Liapunov, Stability of motion: general problem, *Int. J. Control*, 55 (3), Mars 1892 (1992), Lyapunov Centenary Issue.
20. H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*, Springer-Verlag, 1992.
21. W. Perruquetti, *Sur la Stabilité et l'Estimation Des Comportements Non Linéaires, Non Stationnaires, Perturbés*, Ph.D. thesis, University of Sciences and Technology of Lille, France, 1994.
22. H. Reinhard, *Equations Différentielles, Fondements et Applications*, Gauthier-Villars, 1982.
23. J.P. Richard, Edt., *Mathématiques pour les systèmes dynamiques*, Collection I2C, Hermes, Lavoisier, 2002.
24. N. Rouche and J. Mawhin, *Equations Différentielles Ordinaires, Tome 1: Théorie Générale*, Masson et Cie, Paris, 1973.
25. N. Rouche and J. Mawhin, *Equations Différentielles Ordinaires, Tome 2: Stabilité et Solutions Périodiques*, Masson et Cie, Paris, 1973.
26. R. Seydel, *Practical Bifurcation and Stability Analysis: From Equilibrium to Chaos*, 2nd ed., vol. 5 of *IAM*, Springer-Verlag, 1994.

3

Normal Forms and Bifurcations of Vector Fields

C. Dang Vu-Delcarte

CONTENTS

3.1	Introduction	101
3.2	Normal Forms	102
3.2.1	Local Study of the Center Manifold	102
3.2.2	Normal Form Theorem	103
3.2.3	Use of the Basis of Operators $\partial/\partial x_i$	108
3.2.4	Use of Complex Coordinates	111
3.2.5	Resonance	115
3.3	Bifurcations	118
3.3.1	The Hopf Bifurcation	118
3.3.2	Three-Dimensional Systems	123
3.3.2.1	Bifurcation Conditions and Determination of $\alpha'(v_c)$	124
3.3.2.2	Reduction to the Normal Form	125
3.3.2.3	Computation of the Coefficient a_1	126
3.3.3	The Bifurcation in the Rössler System	129
3.3.3.1	Reduction to the Normal Form	130
3.3.3.2	Bifurcation Diagram	133
	References	136

3.1 Introduction

A normal form is the simplest representation of a class of equations featuring a specific bifurcation phenomenon. The normal form is a sufficient

information to understand the dynamical behavior in the neighborhood of a bifurcation. This chapter consists of two parts. In the first part, we describe some of the techniques used to calculate normal forms. The second part deals with applications of the Hopf bifurcation and presents an example of codimension 2 bifurcation.

3.2 Normal Forms

3.2.1 Local Study of the Center Manifold

Let

$$\dot{x} = f(x, \mu), \quad x \in \mathbb{R}^{n+m}, \quad \mu \in \mathbb{R}^k, \quad (\cdot) \equiv \frac{d}{dt} \quad (3.1)$$

be a system of differential equations depending on the k -dimensional parameter μ . Recall (see Chapter 2) that when f is sufficiently smooth, a *fixed point* (or *equilibrium point*) of (3.1) is a point $\bar{x} \in \mathbb{R}^{n+m}$ such that $f(\bar{x}, \mu) = 0$.

Suppose that, by an appropriate change of coordinates, the fixed point $\bar{x}$ has been shifted to the origin. After the transformation, the system of equation reads:

$$\dot{x} = Ax + f(x, y) \quad (3.2a)$$

$$\dot{y} = By + g(x, y) \quad (3.2b)$$

where x and f are n -vectors and A is an $n \times n$ matrix whose eigenvalues have a zero real part; y and g are m -vectors, and B is an $m \times m$ matrix whose eigenvalues have a negative real part (for the sake of simplicity, we will omit the parameters in the right hand side of (3.2) and we assume that the linearized system does not have eigenvalues with a positive real part, namely, $W^u = \emptyset$). In practice, when the Jacobian matrix is diagonalizable (or has Jordan blocks), the dynamical system can be written as in (3.2). This is achieved by using the eigenvector basis. The center manifold, W^c , may be locally represented in the neighborhood of $\bar{x} = 0$ by:

$$W^c(0) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m | y = h(x), |x| < \delta, h(0) = 0, Dh(0) = 0\} \quad (3.3)$$

where $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is defined on some neighborhood $|x| < \delta$ of the origin. Conditions $h(0) = 0$ and $Dh(0) = 0$ imply that $W^c(0)$ is tangent to the center eigenspace $E^c \equiv (y = 0)$ at $(x, y) = (0, 0)$.

Setting $y = h(x)$ in (3.2a), we obtain:

$$\dot{x} = Ax + f(x, h(x)), \quad x \in \mathbb{R}^n, \quad h: \mathbb{R}^n \rightarrow \mathbb{R}^m \quad (3.4)$$

If the fixed point of (3.4) is stable (resp. asymptotically stable), then the fixed point of (3.1) is also stable (resp. asymptotically stable). The nature of the nonhyperbolic fixed point of (3.1) is obtained by looking at the motion on the center manifold. This section describes the theory of normal forms, which uses a coordinate transformation to reduce the system (3.4) to a simpler form containing all the dynamics. The reduced system is called the normal form.

3.2.2 Normal Form Theorem

Let us rewrite (3.4) as:

$$\dot{x} = Ax + F(x), \quad \text{with } F(x) = f(x, h(x)) \quad x \in \mathbb{R}^n \quad (3.5)$$

Let H_k be the vector space spanned by the following vectors

$$x^k e_i \equiv x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} e_i, \quad 1 \leq i \leq n, \quad k_1 + k_2 + \dots + k_n = k \quad (3.6)$$

where $\{e_1, e_2, \dots, e_n\}$ is the basis of the coordinate system $(x_1, x_2, \dots, x_n)$. Let us now perform a Taylor expansion of (3.5) in n variables about the origin:

$$\dot{x} = Ax + F^{(2)}(x) + F^{(3)}(x) + \dots + F^{(k)}(x) + O(|x|^{k+1}). \quad (3.7)$$

$F^{(k)}(x)$ takes the explicit form:

$$F^{(k)}(x) = \left(F_1^{(k)}, F_2^{(k)}, \dots, F_n^{(k)} \right)^T \quad (3.8)$$

where $F_1^{(k)}, \dots, F_n^{(k)}$ are homogeneous polynomials of order k in x . Let us set $L = Ax$. L induces an endomorphism, $\text{ad } L: H_k \rightarrow H_k$, defined by:

$$\text{ad } L(Y) = (DL)Y - (DY)L, \quad \text{for all } Y(x) \in H_k \quad (3.9)$$

where $DL = A$. In the system of coordinates $(x_1, x_2, \dots, x_n)$, (3.9) reads:

$$\text{ad } L(Y)_i = \sum_{j=1}^n \left(\frac{\partial L_i}{\partial x_j} Y_j - \frac{\partial Y_i}{\partial x_j} L_j \right) \quad (3.10)$$

with $L_i = \sum_{j=1}^n A_{ij}x_j$. Let G_k denote the complementary subspace to $\text{ad } L(H_k)$ in H_k (viz., $H_k = \text{ad } L(H_k) \oplus G_k$). The normal form theorem can now be stated:

THEOREM 1

There exists a series of changes of coordinates of the form:

$$x = y + P(y), \quad P(y) \in H_r, \quad r = 2, 3, \dots, k \quad (3.11)$$

that transform the system (3.7) into the normal form:

$$\dot{y} = Ay + g^{(2)}(y) + g^{(3)}(y) + \dots + g^{(k)}(y) + O(|y|^{k+1}) \quad (3.12)$$

where $g^{(i)} \in G_i$, $2 \leq i \leq k$.

This theorem is also known as the Poincaré–Dulac theorem. $P(y)$ is explicitly of the form:

$$P(y) = (P_1, P_2, \dots, P_n)^T \quad (3.13)$$

where $P_1, \dots, P_n$ are homogeneous polynomials of degree r in y .

PROOF The proof consists in the construction of the n -vector $P(y)$ by recurrence. Suppose that we have already performed $k-1$ changes of variables, and that Equation (3.7) at step $k-1$ reads:

$$\dot{x} = Ax + g^{(2)}(x) + g^{(3)}(x) + \dots + g^{(k-1)}(x) + F^{(k)}(x) + O(|x|^{k+1}) \quad (3.14)$$

with $g^{(i)} \in G_i$ for $2 \leq i \leq k-1$ and $F^{(k)}(x) \in H_k$. Let us introduce the change of coordinates:

$$x = y + P(y), \quad P(y) \in H_k \quad (3.15)$$

in (3.14), it follows that:

$$\begin{aligned} (I + DP(y))\dot{y} &= A(y + P(y)) + g^{(2)}(y) + g^{(3)}(y) + \dots \\ &\quad + g^{(k-1)}(y) + F^{(k)}(y) + O(|y|^{k+1}) \end{aligned}$$

Consequently,

$$\begin{aligned} \dot{y} &= Ay + g^{(2)}(y) + g^{(3)}(y) + \dots + g^{(k-1)}(y) \\ &\quad + F^{(k)}(y) + AP(y) - DP(y)Ay + O(|y|^{k+1}) \end{aligned}$$

Terms of degree lower than k are not modified by this transformation. The term of degree k reads:

$$F^{(k)}(y) + DL(y)P(y) - DP(y)L(y) \equiv F^{(k)}(y) + \text{ad } L(P(y)) \quad (3.16)$$

where $L(y) = Ay$, $DL = A$, and $DP = \|\partial P_i / \partial y_j\|$.

Let us seek the conditions for which (3.16) is zero, namely:

$$\text{ad } L(P(y)) = -F^{(k)}(y) \quad (3.17)$$

Recall that the two terms of (3.17) are homogeneous polynomials of degree k . Let $M^{(k)}$ denote the matrix (or the *representation*) of $\text{ad } L$ in the vector space H_k . There are thus two possibilities (the Fredholm alternative):

- Either the matrix $M^{(k)}$ is invertible and hence (3.17) completely determines $P(y)$ and $F^{(k)}$ is eliminated
- Or $M^{(k)}$ is not invertible and hence if G_k denotes the kernel of $M^{(k)}$, we have $F^{(k)} = g^{(k)} + l^{(k)}$ with $g^{(k)} \in G_k = \text{Ker } M^{(k)}$, $l^{(k)} \in M^{(k)}(H_k) = \text{Im } M^{(k)}$, and $g^{(k)}$ cannot be eliminated. In addition, $P(y)$ is not unique.

■

Example 1

Consider the differential system [10]:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} F_1^{(2)}(x, y) \\ F_2^{(2)}(x, y) \end{pmatrix} + O(3) \quad (3.18)$$

with

$$F_1^{(2)}(x, y) = c_{120}x^2 + c_{111}xy + c_{102}y^2$$

$$F_2^{(2)}(x, y) = c_{220}x^2 + c_{211}xy + c_{202}y^2$$

in the canonical basis

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (3.19)$$

has the eigenvalues: $\lambda_1 = \lambda_2 = 0$. The basis of H_2 is:

$$\left\{ \begin{pmatrix} x^2 \\ 0 \end{pmatrix}, \begin{pmatrix} xy \\ 0 \end{pmatrix}, \begin{pmatrix} y^2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ x^2 \end{pmatrix}, \begin{pmatrix} 0 \\ xy \end{pmatrix}, \begin{pmatrix} 0 \\ y^2 \end{pmatrix} \right\}$$

To compute $\text{ad } L(H_2)$, we calculate the action of $\text{ad } L$ on each vector of the basis of H_2 . By virtue of (3.9) it follows that if

$$L = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}$$

we have:

$$\text{ad } L \begin{pmatrix} x^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x^2 \\ 0 \end{pmatrix} - \begin{pmatrix} 2x & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = -2 \begin{pmatrix} xy \\ 0 \end{pmatrix}$$

$$\text{ad } L \begin{pmatrix} xy \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} xy \\ 0 \end{pmatrix} - \begin{pmatrix} y & x \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = - \begin{pmatrix} y^2 \\ 0 \end{pmatrix}$$

$$\text{ad } L \begin{pmatrix} y^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y^2 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 2y \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\text{ad } L \begin{pmatrix} 0 \\ x^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ x^2 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 2x & 0 \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} x^2 \\ -2xy \end{pmatrix}$$

$$\text{ad } L \begin{pmatrix} 0 \\ xy \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ xy \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ y & x \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} xy \\ -y^2 \end{pmatrix}$$

$$\text{ad } L \begin{pmatrix} 0 \\ y^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ y^2 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 2y \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} y^2 \\ 0 \end{pmatrix}$$

Thus, we obtain a basis of $\text{ad } L(H_2)$:

$$\left\{ \begin{pmatrix} xy \\ 0 \end{pmatrix}, \begin{pmatrix} y^2 \\ 0 \end{pmatrix}, \begin{pmatrix} x^2 \\ -2xy \end{pmatrix}, \begin{pmatrix} xy \\ -y^2 \end{pmatrix} \right\}$$

so that $\dim \text{ad } L(H_2) = 4$. Since $\dim H_2 = 6$ and $H_2 = \text{ad } L(H_2) \oplus G_2$, we get $\dim G_2 = 2$. However, the choice of a basis for G_2 is not unique. In fact, we can rewrite the transformation (3.11) in the form:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} P_1(u, v) \\ P_2(u, v) \end{pmatrix} \quad (3.20)$$

with

$$P_1(u, v) = \alpha_{20}u^2 + \alpha_{11}uv + \alpha_{02}v^2$$

$$P_2(u, v) = \beta_{20}u^2 + \beta_{11}uv + \beta_{02}v^2$$

The left-hand side in (3.16) reads:

$$\begin{pmatrix} (c_{120} + \beta_{20})u^2 + (c_{111} + \beta_{11} - 2\alpha_{20})uv + (c_{102} + \beta_{02} - \alpha_{11})v^2 \\ c_{220}u^2 + (c_{211} - 2\beta_{20})uv + (c_{202} - \beta_{11})v^2 \end{pmatrix} \quad (3.21)$$

Note that (3.21) cannot be completely eliminated. We can only reduce the expression to its simplest form.

- If we choose:

$$2\beta_{20} = c_{211}, \quad \beta_{11} = c_{202}, \quad \beta_{11} - 2\alpha_{20} = -c_{111}, \quad \beta_{02} - \alpha_{11} = -c_{102}$$

(3.12) yields (for $k = 2$):

$$\dot{u} = v + \left(\frac{1}{2}c_{211} + c_{120}\right)u^2 + O(3), \quad \dot{v} = c_{220}u^2 + O(3)$$

This is the Takens normal form [9]. In this case, the basis of G_2 is:

$$\left\{ \begin{pmatrix} x^2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ x^2 \end{pmatrix} \right\}$$

- If we choose:

$$\beta_{20} = -c_{120}, \quad \beta_{11} = c_{202}, \quad \beta_{11} - 2\alpha_{20} = -c_{111}, \quad \beta_{02} - \alpha_{11} = -c_{102}$$

(3.12) yields (for $k = 2$):

$$\dot{u} = v + O(3), \quad \dot{v} = c_{220}u^2 + (c_{211} + 2c_{120})uv + O(3)$$

This is the Bogdanov normal form [1]. In this case, the basis of G_2 is:

$$\left\{ \begin{pmatrix} 0 \\ x^2 \end{pmatrix}, \begin{pmatrix} 0 \\ xy \end{pmatrix} \right\}$$

In both cases, the transformation (3.20) is not unique since β_{02} and α_{02} are arbitrary.

REMARK 1

Note that the change of coordinates

$$x_1 = \frac{y_1}{2a}, \quad x_2 = \frac{y_2}{2a} - \frac{y_1^2}{4a}$$

transforms a Takens normal form:

$$\dot{x}_1 = x_2 + ax_1^2 + O(|x|^3), \quad \dot{x}_2 = bx_1^2 + O(|x|^3)$$

into the Bogdanov form:

$$\dot{y}_1 = y_2 + O(|y|^3), \quad \dot{y}_2 = \frac{b}{2a}y_1^2 + y_1y_2 + O(|y|^3)$$

3.2.3 Use of the Basis of Operators $\partial/\partial x_i$

The center manifold, W^c , is a differentiable manifold of class C^r , $r > 1$ and dimension n . Its tangent space at a point $x \in W^c$ will be spanned by the canonical basis [8]:

$$e_i \equiv \frac{\partial}{\partial x_i}, \quad i = 1, 2, \dots, n$$

In this basis, a (tangent) vector x reads:

$$x = \sum_{i=1}^n x_i \frac{\partial}{\partial x_i}$$

and (3.9) yields:

$$\text{ad } L(Y) = \sum_{i=1}^n \sum_{j=1}^n \left(\frac{\partial L_i}{\partial x_j} Y_j - \frac{\partial Y_i}{\partial x_j} L_j \right) \frac{\partial}{\partial x_i} \quad \text{for all } Y \in H_k \quad (3.22)$$

The use of this notation is illustrated with the following example.

Example 2

Consider the following differential system [4]:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} F_1(x, y) \\ F_2(x, y) \end{pmatrix} \quad (3.23)$$

In this case:

$$A = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \quad (3.24)$$

and

$$L = \omega \left(-y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \right) \quad (3.25)$$

The eigenvalues of A are: $\lambda_{1,2} = \pm i\omega$. The action of $\text{ad } L$ on a vector

$$Y = Y_1 \frac{\partial}{\partial x} + Y_2 \frac{\partial}{\partial y}$$

belonging to H_k yields, using (3.22):

$$\begin{aligned} \text{ad } L(Y) &= \omega \left(-Y_2 + y \frac{\partial Y_1}{\partial x} - x \frac{\partial Y_1}{\partial y} \right) \frac{\partial}{\partial x} \\ &\quad + \omega \left(Y_1 + y \frac{\partial Y_2}{\partial x} - x \frac{\partial Y_2}{\partial y} \right) \frac{\partial}{\partial y} \end{aligned} \quad (3.26)$$

The vector space H_2 has the basis:

$$\left\{ x^2 \frac{\partial}{\partial x}, xy \frac{\partial}{\partial x}, y^2 \frac{\partial}{\partial x}, x^2 \frac{\partial}{\partial y}, xy \frac{\partial}{\partial y}, y^2 \frac{\partial}{\partial y} \right\} \quad (3.27)$$

The action of $\text{ad } L$ on these vectors yields, using (3.26):

$$\begin{aligned} \text{ad } L \left(x^2 \frac{\partial}{\partial x} \right) &= 2\omega xy \frac{\partial}{\partial x} + \omega x^2 \frac{\partial}{\partial y} \\ \text{ad } L \left(xy \frac{\partial}{\partial x} \right) &= \omega (y^2 - x^2) \frac{\partial}{\partial x} + \omega xy \frac{\partial}{\partial y} \\ \text{ad } L \left(y^2 \frac{\partial}{\partial x} \right) &= -2\omega xy \frac{\partial}{\partial x} + \omega y^2 \frac{\partial}{\partial y} \\ \text{ad } L \left(x^2 \frac{\partial}{\partial y} \right) &= -\omega x^2 \frac{\partial}{\partial x} + 2\omega xy \frac{\partial}{\partial y} \\ \text{ad } L \left(xy \frac{\partial}{\partial y} \right) &= -\omega xy \frac{\partial}{\partial x} + \omega (y^2 - x^2) \frac{\partial}{\partial y} \\ \text{ad } L \left(y^2 \frac{\partial}{\partial y} \right) &= -\omega y^2 \frac{\partial}{\partial x} - 2\omega xy \frac{\partial}{\partial y} \end{aligned}$$

The resulting six vectors are linearly independent, since the matrix of $\text{ad } L$ in the basis (3.27):

$$M^{(2)} = \omega \begin{bmatrix} 0 & -1 & 0 & -1 & 0 & 0 \\ 2 & 0 & -2 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 2 & 0 & -2 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

is nonsingular (its determinant is equal to $9\omega^6$). We thus have $\text{ad } L(H_2) = H_2$ and $G_2 = \{0\}$. There is consequently no term of degree 2 in the normal form of (3.23).

Let us now seek the terms of degree 3. The vector space H_3 has the basis:

$$\left\{ x^3 \frac{\partial}{\partial x}, x^2 y \frac{\partial}{\partial x}, xy^2 \frac{\partial}{\partial x}, y^3 \frac{\partial}{\partial x}, x^3 \frac{\partial}{\partial y}, x^2 y \frac{\partial}{\partial y}, xy^2 \frac{\partial}{\partial y}, y^3 \frac{\partial}{\partial y} \right\} \quad (3.28)$$

Using (3.26) it follows that

$$\begin{aligned} \text{ad } L \left(x^3 \frac{\partial}{\partial x} \right) &= 3\omega x^2 y \frac{\partial}{\partial x} + \omega x^3 \frac{\partial}{\partial y} \equiv s_1 \\ \text{ad } L \left(x^2 y \frac{\partial}{\partial x} \right) &= \omega(2xy^2 - x^3) \frac{\partial}{\partial x} + \omega x^2 y \frac{\partial}{\partial y} \equiv s_2 \\ \text{ad } L \left(xy^2 \frac{\partial}{\partial x} \right) &= \omega(y^3 - 2x^2 y) \frac{\partial}{\partial x} + \omega xy^2 \frac{\partial}{\partial y} \equiv s_3 \\ \text{ad } L \left(y^3 \frac{\partial}{\partial x} \right) &= -3\omega xy^2 \frac{\partial}{\partial x} + \omega y^3 \frac{\partial}{\partial y} \equiv s_4 \\ \text{ad } L \left(x^3 \frac{\partial}{\partial y} \right) &= -\omega x^3 \frac{\partial}{\partial x} + 3\omega x^2 y \frac{\partial}{\partial y} \equiv s_5 \\ \text{ad } L \left(x^2 y \frac{\partial}{\partial y} \right) &= -\omega x^2 y \frac{\partial}{\partial x} + \omega(2xy^2 - x^3) \frac{\partial}{\partial y} \equiv s_6 \\ \text{ad } L \left(xy^2 \frac{\partial}{\partial y} \right) &= -\omega xy^2 \frac{\partial}{\partial x} + \omega(y^3 - 2x^2 y) \frac{\partial}{\partial y} \equiv s_7 \\ \text{ad } L \left(y^3 \frac{\partial}{\partial y} \right) &= -\omega y^3 \frac{\partial}{\partial x} - 3\omega xy^2 \frac{\partial}{\partial y} \equiv s_8 \end{aligned}$$

The matrix of $\text{ad } L$ in the basis (3.28) is hence:

$$M^{(3)} = \omega \begin{bmatrix} 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 3 & 0 & -2 & 0 & 0 & -1 & 0 & 0 \\ 0 & 2 & 0 & -3 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

An elementary calculation leads to the conclusion that $s_6 = -s_1 - s_3 - s_8$, $s_7 = s_2 + s_4 - s_5$, and $\{s_1, s_2, s_3, s_4, s_5, s_8\}$ are linearly independent. The complementary space G_3 is thus of dimension 2. As we have seen in the

previous example, the choice of a basis of G_3 is not unique. If we choose the following vectors to make up a basis of G_3 :

$$\left\{ (x^2 + y^2) \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right), \quad (x^2 + y^2) \left(-y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \right) \right\} \quad (3.29)$$

the system (3.23) has the normal form:

$$\begin{aligned} \dot{u} &= -\omega v + (a_1 u - b_1 v)(u^2 + v^2) + O(5) \\ \dot{v} &= \omega u + (a_1 v + b_1 u)(u^2 + v^2) + O(5) \end{aligned} \quad (3.30)$$

or, in polar coordinates $u = r \cos \theta$, $v = r \sin \theta$,

$$\begin{aligned} \dot{r} &= a_1 r^3 + O(5) \\ \dot{\theta} &= \omega + b_1 r^2 + O(4) \end{aligned}$$

This is the normal form of the Hopf bifurcation. The choice of the basis (3.29) is justified by the fact that the operator L of (3.25) is invariant with respect to the rotation group, and that the vectors constituting the basis (3.29) share the same property.

3.2.4 Use of Complex Coordinates

The previous computation is simpler if we resort to complex coordinates. Let us set:

$$z = x + iy, \quad \bar{z} = x - iy \quad (3.31)$$

The inverse transformation of (3.31) is:

$$x = \frac{1}{2}(z + \bar{z}), \quad y = \frac{1}{2i}(z - \bar{z})$$

Equation (3.23) may now be rewritten in the form

$$\dot{z} = i\omega z + F(z, \bar{z}), \quad \dot{\bar{z}} = -i\omega \bar{z} + \bar{F}(z, \bar{z}) \quad (3.32)$$

With respect to the variables $z, \bar{z}$, the canonical basis of the tangent space E^c is:

$$\left\{ \frac{\partial}{\partial z}, \frac{\partial}{\partial \bar{z}} \right\}, \quad \text{with } \frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right); \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right)$$

We have:

$$A = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix} \quad \text{and} \quad L = i\omega \left(\frac{\partial}{\partial z} - \frac{\partial}{\partial \bar{z}} \right) \quad (3.33)$$

The action of L on a vector

$$Y = Y_1 \frac{\partial}{\partial z} + Y_2 \frac{\partial}{\partial \bar{z}}$$

belonging to H_k yields, using (3.22):

$$\begin{aligned} \text{ad } L(Y) &= i\omega \left(Y_1 - z \frac{\partial Y_1}{\partial z} + \bar{z} \frac{\partial Y_1}{\partial \bar{z}} \right) \frac{\partial}{\partial z} \\ &\quad - i\omega \left(Y_2 + z \frac{\partial Y_2}{\partial z} - \bar{z} \frac{\partial Y_2}{\partial \bar{z}} \right) \frac{\partial}{\partial \bar{z}} \end{aligned} \quad (3.34)$$

The vector space H_2 has the basis:

$$\left\{ z^2 \frac{\partial}{\partial z}, z\bar{z} \frac{\partial}{\partial z}, \bar{z}^2 \frac{\partial}{\partial z}, z^2 \frac{\partial}{\partial \bar{z}}, z\bar{z} \frac{\partial}{\partial \bar{z}}, \bar{z}^2 \frac{\partial}{\partial \bar{z}} \right\} \quad (3.35)$$

With respect to the basis (3.35), the matrix of $\text{ad } L$ will be diagonal. In fact, using (3.34) we have

$$\begin{aligned} \text{ad } L \left(z^2 \frac{\partial}{\partial z} \right) &= -i\omega z^2 \frac{\partial}{\partial z}, & \text{ad } L \left(z\bar{z} \frac{\partial}{\partial z} \right) &= i\omega z\bar{z} \frac{\partial}{\partial z} \\ \text{ad } L \left(\bar{z}^2 \frac{\partial}{\partial z} \right) &= 3i\omega \bar{z}^2 \frac{\partial}{\partial z}, & \text{ad } L \left(z^2 \frac{\partial}{\partial \bar{z}} \right) &= -3i\omega z^2 \frac{\partial}{\partial \bar{z}} \\ \text{ad } L \left(z\bar{z} \frac{\partial}{\partial \bar{z}} \right) &= -i\omega z\bar{z} \frac{\partial}{\partial \bar{z}}, & \text{ad } L \left(\bar{z}^2 \frac{\partial}{\partial \bar{z}} \right) &= i\omega \bar{z}^2 \frac{\partial}{\partial \bar{z}} \end{aligned}$$

It follows that the matrix of $\text{ad } L$ in H_2 yields:

$$M^{(2)} = i\omega \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and that $\det M^{(2)} = 9\omega^6$, as in the case of the real variables. Consequently, the terms of degree 2 in (3.32) may be eliminated.

The vector space H_3 has the basis:

$$\left\{ z^3 \frac{\partial}{\partial z}, z^2 \bar{z} \frac{\partial}{\partial z}, z \bar{z}^2 \frac{\partial}{\partial z}, \bar{z}^3 \frac{\partial}{\partial z}, z^3 \frac{\partial}{\partial \bar{z}}, z^2 \bar{z} \frac{\partial}{\partial \bar{z}}, z \bar{z}^2 \frac{\partial}{\partial \bar{z}}, \bar{z}^3 \frac{\partial}{\partial \bar{z}} \right\}$$

The action of $\text{ad } L$ on these vector is, according to (3.34),

$$\begin{aligned} \text{ad } L \left(z^3 \frac{\partial}{\partial z} \right) &= -2i\omega z^3 \frac{\partial}{\partial z}, & \text{ad } L \left(z^2 \bar{z} \frac{\partial}{\partial z} \right) &= 0 \\ \text{ad } L \left(z \bar{z}^2 \frac{\partial}{\partial z} \right) &= 2i\omega \bar{z}^2 \frac{\partial}{\partial z}, & \text{ad } L \left(\bar{z}^3 \frac{\partial}{\partial z} \right) &= 4i\omega \bar{z}^3 \frac{\partial}{\partial z} \\ \text{ad } L \left(z^3 \frac{\partial}{\partial \bar{z}} \right) &= -4i\omega z^3 \frac{\partial}{\partial \bar{z}}, & \text{ad } L \left(z^2 \bar{z} \frac{\partial}{\partial \bar{z}} \right) &= -2i\omega z^2 \bar{z} \frac{\partial}{\partial \bar{z}} \\ \text{ad } L \left(z \bar{z}^2 \frac{\partial}{\partial \bar{z}} \right) &= 0, & \text{ad } L \left(\bar{z}^3 \frac{\partial}{\partial \bar{z}} \right) &= 2i\omega \bar{z}^3 \frac{\partial}{\partial \bar{z}} \end{aligned}$$

We immediately remark that the matrix $M^{(3)}$

$$M^{(3)} = i\omega \begin{bmatrix} -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

is not invertible and that $G_3 \equiv \text{Ker } M^{(3)}$ is of dimension 2 and is spanned by the following vectors

$$\left\{ z^2 \bar{z} \frac{\partial}{\partial z}, z \bar{z}^2 \frac{\partial}{\partial \bar{z}} \right\}$$

More generally, it follows from (3.34) that:

$$\begin{aligned} \text{ad } L \left(z^k \bar{z}^l \frac{\partial}{\partial z} \right) &= i\omega(1 - k + l) z^k \bar{z}^l \frac{\partial}{\partial z} \\ \text{ad } L \left(z^k \bar{z}^l \frac{\partial}{\partial \bar{z}} \right) &= -i\omega(1 + k - l) z^k \bar{z}^l \frac{\partial}{\partial \bar{z}} \end{aligned}$$

Therefore, we have:

$$\text{ad } L \left(z^{l+1} \bar{z}^l \frac{\partial}{\partial z} \right) = \text{ad } L \left(z^l \bar{z}^{l+1} \frac{\partial}{\partial \bar{z}} \right) = 0$$

and the matrix $M^{(2l+1)}$ will be noninvertible. $G_{2l+1} = \text{Ker } M^{(2l+1)}$ is spanned by the following vectors

$$\left\{ z^{l+1} \bar{z}^l \frac{\partial}{\partial z}, z^l \bar{z}^{l+1} \frac{\partial}{\partial \bar{z}} \right\}$$

Consequently, there exists a transformation:

$$z = \xi + \chi(\xi, \bar{\xi}), \quad \text{degree } \chi(\xi, \bar{\xi}) > 1$$

converting (3.32) into

$$\dot{\xi} = \lambda \xi + c_1 \xi^2 \bar{\xi} + c_2 \xi^3 \bar{\xi}^2 + \dots + c_l \xi^{l+1} \bar{\xi}^l + \dots \quad (3.36)$$

where $\lambda = i\omega$. At the third order, (3.36) is identical to (3.30) with $\xi = u + iv$, $c_1 = a_1 + ib_1$. The normal form (3.36) is known as the Poincaré normal form and plays a fundamental role in the analysis of the Hopf bifurcation (see Section 3.3.1).

Example 3

Consider the system:

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = -x^2 y - x \quad (3.37)$$

(Van der Pol equation with $\varepsilon = 0$) [6]. This system has the form (3.23) with $\omega = -1$. Setting $z = x + iy$, (3.37) yields

$$\frac{dz}{dt} = -iz - \frac{z^3 + z^2 \bar{z} - z \bar{z}^2 - \bar{z}^3}{8} \quad (3.38)$$

To reduce (3.38) to the normal form (3.36), we perform a change of variables of the form:

$$z = \xi + \alpha \xi^3 + \beta \xi^2 \bar{\xi} + \gamma \xi \bar{\xi}^2 + \delta \bar{\xi}^3 \quad (3.39)$$

Replacing (3.39) in (3.38) and neglecting the terms of order ≥ 4 , we get:

$$\begin{aligned} & \left(1 + 3\alpha \xi^2 + 2\beta \xi \bar{\xi} + \gamma \bar{\xi}^2\right) \frac{d\xi}{dt} + \left(\beta \xi^2 + 2\gamma \xi \bar{\xi} + 3\delta \bar{\xi}^2\right) \frac{d\bar{\xi}}{dt} \\ &= -i \left(\xi + \alpha \xi^3 + \beta \xi^2 + \bar{\xi} + \gamma \xi \bar{\xi}^2 + \delta \bar{\xi}^3 \right) \\ & \quad - \frac{\xi^3 + \xi^2 \bar{\xi} - \xi \bar{\xi}^2 - \bar{\xi}^3}{8}. \end{aligned}$$

Having neglected the terms of order greater than four, one can simply replace $d\bar{\xi}/dt$ with $i\bar{\xi}$ and isolate $d\xi/dt$ by multiplying both terms of the equation by $1 - (3\alpha\bar{\xi}^2 + 2\beta\xi\bar{\xi} + \gamma\bar{\xi}^2)$. It follows that:

$$\begin{aligned} \frac{d\xi}{dt} = & -i\xi + \left(2\alpha - \frac{1}{8}\right)\xi^3 - \frac{1}{8}\xi^2\bar{\xi} \\ & - \left(2\gamma - \frac{1}{8}\right)\xi\bar{\xi}^2 - \left(4\delta - \frac{1}{8}\right)\bar{\xi}^3 \end{aligned}$$

As expected, the term in $\xi^2\bar{\xi}$ cannot be eliminated. Nevertheless, we can choose $\alpha = \gamma = 2\delta = 1/16$ in order to eliminate the other terms:

$$\frac{d\xi}{dt} = -i\xi - \frac{1}{8}|\xi|^2\xi$$

The change of variables (3.39) is not unique since β remains arbitrary.

3.2.5 Resonance

We now examine the procedure used to find the normal form of a system of differential equations through the eigenvalues of the Jacobian matrix. We start with a system of differential equations

$$\frac{dx}{dt} = f(x), \quad x \in \mathbb{R}^n, \quad f: \mathbb{R}^n \longrightarrow \mathbb{R}^n \quad (3.40)$$

which has an equilibrium at 0. Let A be the Jacobian matrix of (3.40) at $x = 0$. Suppose that the matrix A has n distinct eigenvalues and let e_i be the eigenvectors corresponding to the eigenvalues λ_i , $i = 1, 2, \dots, n$. In addition, suppose that a (linear) coordinate transformation has been performed so that $(x_1, x_2, \dots, x_n)$ are the coordinates with respect to the eigenbasis $(e_1, e_2, \dots, e_n)$. The matrix A is thus diagonal in this basis. The matrix of $\text{ad } L$ in H_k will also be diagonal, and its eigenvectors are $x^k e_i$ where we have set (see (3.6)):

$$x^k = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}, \quad \text{with } k_1 + k_2 + \dots + k_n = k \geq 2 \quad (3.41)$$

Indeed, (3.9) gives:

$$\text{ad } L(x^k e_i) = (DL)x^k e_i - D(x^k e_i)L \quad (3.42)$$

but

$$D(x^k e_i)Ax = \sum_j \frac{k_j x^k}{x_j} \lambda_j x_j e_i = \sum_j k_j \lambda_j x^k e_i$$

and $Ax^k e_i = \lambda_i x^k e_i$; so that, according to (3.42):

$$\text{ad } L(x^k e_i) = \left(\lambda_i - \sum_{j=1}^n k_j \lambda_j \right) x^k e_i \quad (3.43)$$

$x^k e_i$ is thus an eigenvector of $\text{ad } L$ in H_k corresponding to the eigenvalue $\lambda_i - \sum_{j=1}^n k_j \lambda_j$. If every eigenvalue of $\text{ad } L$ in H_k is non-zero, $\text{ad } L$ is invertible and $G_k = \text{Ker}(\text{ad } L) = \{0\}$.

DEFINITION 1 The eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ are said to be resonant of order k if there exists an eigenvalue λ_i such that:

$$\lambda_i = \sum_{j=1}^n k_j \lambda_j, \quad \sum_{j=1}^n k_j = k \geq 2, \quad k_j \geq 0 \quad (3.44)$$

Relation (3.44) is equivalent to $\lambda_i = (k, \lambda)$. If (3.44) is satisfied, the terms in (3.41) are called *resonant terms*: they cannot be eliminated.

Example 4

The matrix A in (3.24) has eigenvalues $\lambda = i\omega$, $\bar{\lambda} = -i\omega$. In the basis of eigenvectors

$$e_1 = \frac{1}{2} \begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad e_2 = \frac{1}{2} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

A takes the form:

$$P^{-1}AP = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix} \quad \text{with } P = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

and the variables, in the eigenvector basis, are $z, \bar{z}$ (with $z = x + iy$, $\bar{z} = x - iy$). There are resonances of order $2l + 1$, $l \geq 1$, since we can write:

$$\lambda = (l + 1)\lambda + l\bar{\lambda}$$

or, taking the complex conjugate,

$$\bar{\lambda} = (l + 1)\bar{\lambda} + l\lambda$$

There are no resonances of order $2l$. Therefore, for an appropriate change of variables:

$$z = \xi + \chi(\xi, \bar{\xi}), \quad \text{degree } \chi(\xi, \bar{\xi}) > 1$$

the normal form will read:

$$\dot{\xi} = \lambda \xi + c_1 \xi^2 \bar{\xi} + c_2 \xi^3 \bar{\xi}^2 + \cdots + c_l \xi^{l+1} \bar{\xi}^l + \cdots$$

and we encounter once again the Poincaré normal form (see (3.36)).

Example 5

Consider the Lorenz system:

$$\begin{aligned} \frac{dx}{dt} &= -\sigma x + \sigma y \\ \frac{dy}{dt} &= -xz + rx - y \\ \frac{dz}{dt} &= xy - bz \end{aligned} \tag{3.45}$$

The system has a couple of nontrivial fixed points

$$x^* = y^* = \pm \sqrt{b(r-1)}, \quad z^* = r-1$$

The Hopf bifurcation takes place at (see Section 3.3.1)

$$r = r_c = \frac{\sigma(\sigma + b + 3)}{\sigma - b - 1}$$

Perform the following change of variables:

$$u = x - x^*, \quad v = y - y^*, \quad w = z - z^*$$

to shift the fixed point to the origin. We obtain the system:

$$\begin{aligned} \begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{bmatrix} &= \begin{bmatrix} -\sigma & \sigma & 0 \\ 1 & -1 & -x^* \\ y^* & x^* & -b \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} + \begin{bmatrix} 0 \\ -uw \\ uv \end{bmatrix}, \\ x^* &= \pm \sqrt{b(r_c - 1)} \end{aligned} \tag{3.46}$$

The eigenvalues of the matrix in (3.46) are given by (see (3.87))

$$\lambda_{1,2} = \pm i\omega_0, \quad \text{with } \omega_0 = \sqrt{b(r_c + \sigma)} \quad \text{and } \lambda_3 = -(\sigma + b + 1)$$

The resonances

$$\lambda_i = k_1 \lambda_1 + k_2 \lambda_2 + k_3 \lambda_3 \equiv (k, \lambda)$$

are:

- For $i = 1$: $(k_1, k_2, k_3) = (n + 1, n, 0)$
- For $i = 2$: $(k_1, k_2, k_3) = (n, n + 1, 0)$
- For $i = 3$: $(k_1, k_2, k_3) = (n, n, 1) \quad n = 1, 2, 3, \dots$

The normal form of the Lorenz system at $r = r_c$ is hence:

$$\begin{aligned}\dot{y}_1 &= \lambda_1 y_1 + c_1 y_2 y_1^2 + \dots + c_n y_2^n y_1^{n+1}, & y_2 &= \bar{y}_1, \\ \dot{y}_3 &= \lambda_3 y_3 + d_1 (y_1 y_2) y_3 + \dots + d_n (y_1 y_2)^n y_3\end{aligned}\tag{3.47}$$

where $c_n \in \mathbb{C}$, $d_n \in \mathbb{R}$. In the next section (see Example 6) we will establish how the Lorenz system (3.46) can be taken to its normal form (3.47).

From the normal form theorem, Theorem 2 follows.

THEOREM 2

If the eigenvalues of A are nonresonant then, the equation

$$\dot{x} = Ax + F(x), \quad \text{degree } F(x) > 1\tag{3.48}$$

may be reduced to a linear equation $\dot{y} = Ay$, through a change of variables

$$x = y + P(y)$$

3.3 Bifurcations

In this section, as applications of the normal forms theory, we consider the Hopf bifurcation and the bifurcation of the Rössler system.

3.3.1 The Hopf Bifurcation

Suppose that the dynamical system governed by the equation

$$\dot{u} = f(u, v), \quad u \in \mathbb{R}^n, \quad v: \text{real parameter}\tag{3.49}$$

has an equilibrium point $u = u^*(v)$ and that

(H): the Jacobian matrix

$$A(v) = \left\| \frac{\partial f_i}{\partial u_j} \right\|_{u=u^*}$$

has a couple of complex conjugate eigenvalues λ_1 and λ_2 ,

$$\lambda_{1,2}(v) = \alpha(v) \pm i\omega(v)$$

such that:

1. For a certain value $v = v_c$

$$\alpha(v_c) = 0 \quad \text{and} \quad \left. \frac{d}{dv} \alpha(v) \right|_{v=v_c} \neq 0$$

2. The $n - 2$ remaining eigenvalues of $A(v_c)$ have a strictly negative real part.

At the point $u = u^*(v_c)$, we have a two-dimensional center manifold and a stable manifold of dimension $n - 2$. Perform a coordinate transformation so that (3.49) can be written in the form (3.2). We start by replacing

$$u \longrightarrow u^* + u, \quad v = v_c + \mu \quad (3.50)$$

in (3.49) so that the fixed point is shifted to the origin, and so that the value v_c is shifted to 0. Equation (3.49) may now be written in the form:

$$\dot{u} = A(\mu)u + \hat{F}(u, \mu) \quad (3.51)$$

where $\hat{F}(u, \mu)$ is the nonlinear term.

Let $v_1(\mu)$ (resp. $v_2(\mu) = \bar{v}_1$) be the eigenvector of $A(\mu)$ corresponding to the eigenvalue $\lambda_1(\mu) = \alpha(\mu) + i\omega(\mu)$ (resp. $\lambda_2(\mu) = \alpha(\mu) - i\omega(\mu)$). Consider the following basis $\{e_1, e_2, \dots, e_n\}$ where $e_1 = \Re v_1$, $e_2 = -\Im v_1$, and $\{e_3, \dots, e_n\}$ is a *real* basis of the union of the eigenspaces of $\lambda_3, \dots, \lambda_n$. Let T be the transformation matrix whose columns are $\{e_1, e_2, \dots, e_n\}$:

$$T = [e_1 \ e_2 \ \dots \ e_n] \quad (3.52)$$

Replacing the change of variables,

$$u = Tx, \quad x = T^{-1}u \quad (3.53)$$

in (3.51), we get:

$$\dot{x} = A'(\mu)x + F(x, \mu) \quad (3.54)$$

with

$$A'(\mu) = T^{-1}A(\mu)T = \begin{bmatrix} \alpha(\mu) & -\omega(\mu) & 0 \\ \omega(\mu) & \alpha(\mu) & 0 \\ 0 & 0 & B(\mu) \end{bmatrix} \quad (3.55)$$

where $B(\mu)$ is an $(n-2) \times (n-2)$ matrix and

$$F(x, \mu) = T^{-1}\hat{F}(Tx, \mu) \quad (3.56)$$

Let us set:

$$\begin{aligned} z &= x_1 + ix_2, \\ y &= (x_3, x_4, \dots, x_n)^T \end{aligned} \quad (3.57)$$

Hence, (3.54) reads:

$$\dot{z} = \lambda(\mu)z + G(z, \bar{z}, y, \mu), \quad \lambda(\mu) = \alpha(\mu) + i\omega(\mu) \quad (3.58)$$

$$\dot{y} = B(\mu)y + H(z, \bar{z}, y) \quad (3.59)$$

where we have set

$$G(z, \bar{z}, y, \mu) = F^1(x_1, x_2, y, \mu) + iF^2(x_1, x_2, y, \mu) \quad (3.60)$$

Let

$$y = w(z, \bar{z}) \quad (3.61)$$

be the center manifold equation; then, the next step in the procedure is to transform (3.58) into its Poincaré normal form (3.36):

$$\dot{\xi} = \lambda(\mu)\xi + c_1(\mu)\bar{\xi}\xi^2 + \dots + c_k(\mu)\bar{\xi}^k\xi^{k+1} + \dots, \quad c_k(\mu) \in \mathbb{C} \quad (3.62)$$

by means of a transformation of the type:

$$z = \xi + \chi(\xi, \bar{\xi})$$

Equation (3.62) in polar coordinates ($\xi = r e^{i\theta}$), reads:

$$\begin{aligned} \dot{r} &= r[\alpha(\mu) + a_1r^2 + \dots] \\ \dot{\theta} &= \omega(\mu) + b_1r^2 + \dots \end{aligned} \quad (3.63)$$

where we have set: $a_i = \Re c_i$ and $b_i = \Im c_i$. At first order, we have $\omega(\mu) = \omega(0) + \dots$ and $\alpha(\mu) = \alpha'(0)\mu + \dots$. It follows that $r = \text{const.}$ if $\alpha'(0)\mu + a_1r^2 = 0$, and thus we have the following theorem [5, 7].

THEOREM 3

If hypotheses (H) are satisfied and if $a_1 \neq 0$, $\alpha'(0)\mu/a_1 < 0$, then the fixed point $u^*(v_c)$ bifurcates into a limit cycle of radius

$$r \approx \sqrt{-\frac{\alpha'(0)\mu}{a_1}}, \quad \dot{r} = \frac{d}{d\mu} \quad (3.64)$$

and of period $T \approx 2\pi/\omega_0$ with $\omega_0 = \omega(0)$.

Let us set:

$$\delta = -\frac{a_1}{\alpha'(0)} \quad (3.65)$$

Since r must be a positive real value, (3.64) shows that the periodic orbits appear (or the direction of the bifurcation is) on the side $\mu < 0$ if $\delta < 0$ and that the periodic orbits appear (or the direction of the bifurcation is) on the side $\mu > 0$ if $\delta > 0$. If $a_1 = 0$, we must perform an expansion of order > 1 .

It follows from (3.63) that at first order we have

$$\frac{dr}{d\theta} = \frac{r}{\omega_0} [\alpha'(0)\mu + a_1 r^2] + \dots \quad (3.66)$$

and hence the stability of the fixed point is determined by the sign of $\alpha'(0)\mu$. There are now four possibilities [10]:

- *Case 1:* $\alpha'(v_c) > 0$ and $a_1 > 0$. In this case, the origin is an unstable fixed point if $\mu > 0$ and an asymptotically stable fixed point if $\mu < 0$, with an unstable periodic orbit if $\mu < 0$ (there is no periodic orbit if $\mu > 0$) (see Figure 3.1a).
- *Case 2:* $\alpha'(v_c) > 0$ and $a_1 < 0$. In this case, the origin is an asymptotically stable fixed point if $\mu < 0$ and an unstable fixed point if $\mu > 0$, with an asymptotically stable periodic orbit if $\mu > 0$ (there is no periodic orbit if $\mu < 0$) (see Figure 3.1b).
- *Case 3:* $\alpha'(v_c) < 0$ and $a_1 > 0$. In this case, the origin is an unstable fixed point if $\mu < 0$ and an asymptotically stable fixed point if $\mu > 0$, with an unstable periodic orbit if $\mu > 0$ (there is no periodic orbit if $\mu < 0$) (see Figure 3.1c).
- *Case 4:* $\alpha'(v_c) < 0$ and $a_1 < 0$. In this case, the origin is an asymptotically stable fixed point if $\mu > 0$ and an unstable fixed point if $\mu < 0$, with an asymptotically stable periodic orbit if $\mu < 0$ (there is no periodic orbit if $\mu > 0$) (see Figure 3.1d).

Take for instance Case 4 with $\mu < 0$. For a small enough μ , consider the annulus A defined by (see Figure 3.2):

$$A = \{(r, \theta) | r_1 \leq r \leq r_2\}$$

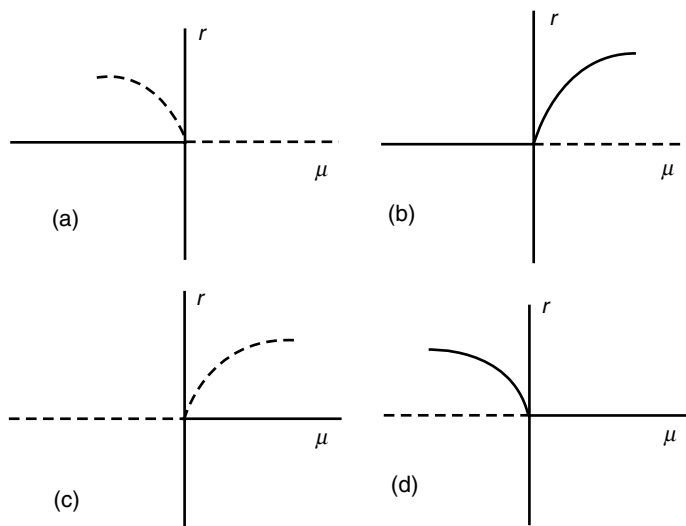


FIGURE 3.1
Hopf bifurcation diagrams.

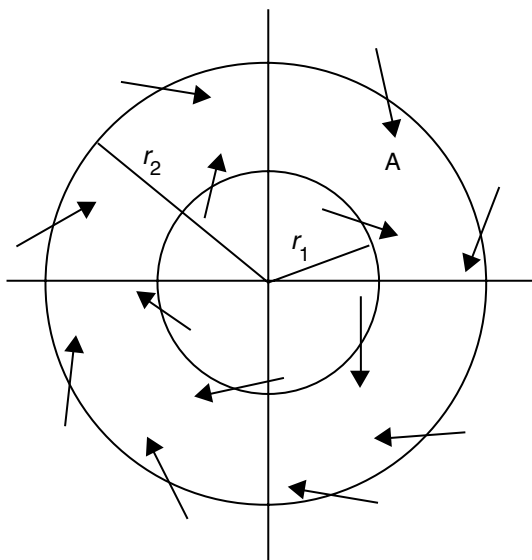
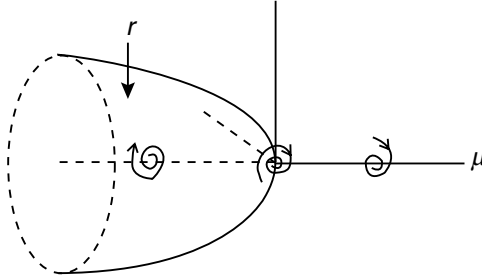


FIGURE 3.2
The Poincaré-Bendixson annulus.

**FIGURE 3.3**

$\alpha'(v_c) < 0$ and $a_1 < 0$.

where r_1 and r_2 are chosen so that

$$0 < r_1 < \sqrt{\frac{-\alpha'(0)\mu}{a_1}} < r_2$$

We can easily check that the right-hand side in (3.66) is positive when $r = r_1$ and negative for $r = r_2$. It follows that, along the boundary of A , the vector field is directed everywhere inward; so, by virtue of the Poincaré–Bendixson theorem, A contains a stable periodic orbit. The remaining cases are proved likewise.

It follows that if $a_1 < 0$ (resp. $a_1 > 0$), the periodic orbit will be asymptotically stable (resp. unstable) and the Hopf bifurcation is said to be *supercritical* (resp. *subcritical*). The coefficient a_1 is called a *Liapunov number* [10].

3.3.2 Three-Dimensional Systems

Consider a three dimensional system depending on a single parameter

$$\dot{u} = f(u, v), \quad u \in \mathbb{R}^3 \quad (3.67)$$

where, for the sake of simplicity, we have assumed that the fixed point has been shifted to the origin (by a suitable coordinate transformation). We may thus write (3.67) in the form:

$$\dot{u} = A(v)u + \hat{F}(u, v) \quad \text{with } A(v) \equiv \left\| \frac{\partial f_i}{\partial u_j}(v) \right\|_{u=0} \quad (3.68)$$

3.3.2.1 Bifurcation Conditions and Determination of $\alpha'(v_c)$

The characteristic equation of the matrix A reads:

$$|a_{ij}(v) - \lambda \delta_{ij}| = 0$$

or

$$\lambda^3 + P(v)\lambda^2 + Q(v)\lambda + R(v) = 0 \quad (3.69)$$

with

$$P(v) = -\text{Tr } A = -\sum_{i=1}^3 a_{ii}(v)$$

$$Q(v) = \text{Tr } A^c = \begin{vmatrix} a_{11}(v) & a_{12}(v) \\ a_{21}(v) & a_{22}(v) \end{vmatrix} + \begin{vmatrix} a_{11}(v) & a_{13}(v) \\ a_{31}(v) & a_{33}(v) \end{vmatrix} + \begin{vmatrix} a_{22}(v) & a_{23}(v) \\ a_{32}(v) & a_{33}(v) \end{vmatrix}$$

$$R(v) = -\det |a_{ij}(v)|$$

Suppose that (3.69) has a couple of complex conjugate roots $\lambda_{1,2}(v) = \alpha(v) \pm i\omega(v)$ and a real root $\lambda_3(v)$ such that, for a certain value $v = v_c$, we have:

$$\alpha(v_c) = 0, \quad \alpha'(v_c) \neq 0, \quad \lambda_3(v_c) < 0$$

where we have set $(') = d/dv$.

According to the relations between roots and coefficients of a polynomial, we have:

$$2\alpha(v) + \lambda_3(v) = -P(v)$$

$$\alpha(v)^2 + \omega(v)^2 + 2\alpha(v)\lambda_3(v) = Q(v)$$

$$[\alpha(v)^2 + \omega(v)^2]\lambda_3(v) = -R(v)$$

It follows that for $v = v_c$:

$$\lambda_3(v_c) = -P(v_c), \quad \omega(v_c)^2 = Q(v_c), \quad \omega(v_c)^2 \lambda_3(v_c) = -R(v_c) \quad (3.70)$$

Thus, the conditions for the coefficients $P(v)$, $Q(v)$, and $R(v)$ are, respectively:

$$P(v_c) > 0, \quad Q(v_c) > 0, \quad P(v_c)Q(v_c) = R(v_c). \quad (3.71)$$

This equation allows us to compute the bifurcation point v_c (if it exists). Differentiate (3.69) with respect to v :

$$3\lambda^2\lambda' + P'\lambda^2 + 2P\lambda\lambda' + Q'\lambda + Q\lambda' + R' = 0 \quad (3.72)$$

and replace $\lambda = \alpha(v) + i\omega(v)$ and $\lambda' = \alpha'(v) + i\omega'(v)$ in (3.72). With $v = v_c$, we get:

$$\begin{aligned} 2Q(v_c)\alpha'(v_c) &= R'(v_c) - P'(v_c)Q(v_c) - 2P(v_c)\omega(v_c)\omega'(v_c) \\ 2\omega(v_c)\omega'(v_c) &= 2\alpha'(v_c)P(v_c) + Q'(v_c) \end{aligned}$$

It follows that:

$$\alpha'(v_c) = \frac{R'(v_c) - P'(v_c)Q(v_c) - P(v_c)Q'(v_c)}{2[Q(v_c) + P^2(v_c)]} \quad (3.73)$$

and thus:

$$\text{sgn} [\alpha'(v_c)] = \text{sgn} [R'(v_c) - P'(v_c)Q(v_c) - P(v_c)Q'(v_c)] \quad (3.74)$$

3.3.2.2 Reduction to the Normal Form

Let

$$v_i = (\alpha_i, \beta_i, \gamma_i), \quad i = 1, 2, 3$$

be the eigenvector of the matrix $A = \|a_{ij}(v_c)\|$ corresponding to the eigenvalue λ_i , with

$$\lambda_{1,2} = \pm i\omega_0 = \pm iQ^{1/2}(v_c), \quad \lambda_3 = -P(v_c) \quad (3.75)$$

Hence:

$$(A - \lambda_i I)v_i = 0 \quad (3.76)$$

The solution to the homogeneous equation (3.76) depends on an arbitrary constant. We could, for instance, compute β_i, γ_i as a function of α_i . If we replace the solutions v_i of (3.76) in

$$T \equiv (t_{ij}) = [\Re v_1 \quad -\Im v_1 \quad v_3] \quad (3.77)$$

we obtain the following expressions for the elements t_{ij} of the matrix T :

$$\begin{aligned}
 t_{11} &= \alpha_1, \quad t_{12} = 0, \quad t_{13} = \alpha_3 \\
 t_{21} &= \frac{\alpha_1 [(a_{21}a_{13} - a_{23}a_{11})\Delta + a_{13}a_{23}\omega_0^2]}{D_1} \\
 t_{22} &= \frac{-\alpha_1\omega_0[\Delta a_{23} - (a_{21}a_{13} - a_{23}a_{11})a_{13}]}{D_1} \\
 t_{23} &= \frac{\alpha_3[a_{23}\lambda_3 + a_{13}a_{21} - a_{11}a_{23}]}{D_2} \\
 t_{31} &= \frac{\alpha_1 [(a_{11}a_{22} - a_{12}a_{21} - \omega_0^2)\Delta - a_{13}(a_{11} + a_{22})\omega_0^2]}{D_1} \\
 t_{32} &= \frac{\alpha_1\omega_0 [a_{13}(a_{11}a_{22} - a_{12}a_{21} - \omega_0^2) + (a_{11} + a_{22})\Delta]}{D_1} \\
 t_{33} &= \frac{\alpha_3[(a_{11} - \lambda_3)(a_{22} - \lambda_3) - a_{12}a_{21}]}{D_2}
 \end{aligned} \tag{3.78}$$

where we have set:

$$\Delta = a_{12}a_{23} - a_{13}a_{22}, \quad D_1 = \Delta^2 + a_{13}^2\omega_0^2, \quad D_2 = \Delta + a_{13}\lambda_3$$

where α_1 and α_3 are two arbitrary real numbers (for the sake of simplicity, we have assumed that $\Im\alpha_1 = 0$).

3.3.2.3 Computation of the Coefficient a_1

The system (3.54) at $v = v_c$ reads:

$$\dot{x} = A'x + F(x) \tag{3.79}$$

where $u = Tx$, $x = T^{-1}u$ and

$$A' = T^{-1}AT = \begin{bmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad F(x) = T^{-1}\hat{F}(Tx, v_c) \tag{3.80}$$

The calculation sketched in [3] yields:

$$\begin{aligned}
 16a_1 = & \frac{1}{\omega_0} \left[\left(F_{x_1x_1}^1 + F_{x_2x_2}^1 \right) F_{x_1x_2}^1 \right. \\
 & - \left(F_{x_1x_1}^2 + F_{x_2x_2}^2 \right) F_{x_1x_2}^2 - F_{x_1x_1}^1 F_{x_1x_1}^2 + F_{x_2x_2}^1 F_{x_2x_2}^2 \left. \right] \\
 & + \left(F_{x_1x_1x_1}^1 + F_{x_1x_2x_2}^1 + F_{x_1x_1x_2}^2 + F_{x_2x_2x_2}^2 \right) \\
 & - \frac{2}{\lambda_3} \left(F_{x_1x_3}^1 + F_{x_2x_3}^2 \right) \left(F_{x_1x_1}^3 + F_{x_2x_2}^3 \right) \\
 & - \frac{1}{4\omega_0^2 + \lambda_3^2} \left(F_{x_1x_3}^1 - F_{x_2x_3}^2 \right) \left[\lambda_3 \left(F_{x_1x_1}^3 - F_{x_2x_2}^3 \right) + 4\omega_0 F_{x_1x_2}^3 \right] \\
 & + \frac{2}{4\omega_0^2 + \lambda_3^2} \left(F_{x_2x_3}^1 + F_{x_1x_3}^2 \right) \left[\omega_0 \left(F_{x_1x_1}^3 - F_{x_2x_2}^3 \right) - \lambda_3 F_{x_1x_2}^3 \right] \quad (3.81)
 \end{aligned}$$

The indices denote partial derivatives. ω_0 and λ_3 are given by (3.75).

Example 6

Let us determine the nature of the Hopf bifurcation for the Lorenz system:

$$\begin{aligned}
 \dot{u}_1 &= -\sigma u_1 + \sigma u_2 \\
 \dot{u}_2 &= -u_1 u_3 + r u_1 - u_2 \\
 \dot{u}_3 &= u_1 u_2 - b u_3
 \end{aligned} \quad (3.82)$$

Recall that at $r > 1$ (see Example 5) the system has two fixed points C and C' located, respectively, at:

$$u_1^* = u_2^* = \pm \sqrt{b(r-1)}, \quad u_3^* = r-1$$

Let us replace $u \rightarrow u^* + u$ in (3.82) in order to take (3.82) to the form (3.68):

$$\dot{u} = Au + \hat{F}(u) \quad (3.83)$$

we get:

$$A = \left\| \frac{\partial f_i}{\partial u_j} \right\|_{u=u^*} = \begin{bmatrix} -\sigma & \sigma & 0 \\ 1 & -1 & -u_1^* \\ u_2^* & u_1^* & -b \end{bmatrix}$$

$$\hat{F}^1(u) = 0, \quad \hat{F}^2(u) = -u_1 u_3, \quad \hat{F}^3(u) = u_1 u_2 \quad (3.84)$$

The matrix A has the characteristic polynomial:

$$\lambda^3 + (\sigma + b + 1)\lambda^2 + b(r + \sigma)\lambda + 2b\sigma(r - 1) = 0 \quad (3.85)$$

For

$$r = r_c = \frac{\sigma(\sigma + b + 3)}{\sigma - b - 1} \quad (3.86)$$

Equation (3.85) has two pure imaginary roots and a negative real root:

$$\lambda_{1,2} = \pm i\omega_0, \quad \text{with } \omega_0 = \sqrt{b(r_c + \sigma)} \text{ and } \lambda_3 = -(\sigma + b + 1) \quad (3.87)$$

Thus, replacing $P(r_c) = \sigma + b + 1 = -\lambda_3$, $P'(r_c) = 0$, $Q(r_c) = b(r_c + \sigma) = \omega_0^2$, $Q'(r_c) = b$, $R(r_c) = 2b\sigma(r_c - 1)$, $R'(r_c) = 2b\sigma$ into (3.73), we obtain:

$$\alpha'(r_c) = \frac{b(\sigma - b - 1)}{2(\omega_0^2 + \lambda_3^2)} \quad (3.88)$$

With the choice $\alpha_1 = 1$, $\alpha_3 = 1$, in (3.78) we get:

$$T = \frac{1}{\sigma u_1^*} \begin{bmatrix} \sigma u_1^* & 0 & \sigma u_1^* \\ \sigma u_1^* & -\omega_0 u_1^* & (\sigma + \lambda_3)u_1^* \\ \omega_0^2 & \omega_0(1 + \sigma) & b\lambda_3 \end{bmatrix} \equiv \|t_{ij}\| \quad (3.89)$$

It follows that:

$$T^{-1} = \frac{1}{\omega_0(\lambda_3^2 + \omega_0^2)} \begin{bmatrix} -\omega_0[b\lambda_3 + (\sigma + \lambda_3)(1 + \sigma)] & \omega_0\sigma(1 + \sigma) & \sigma\omega_0 u_1^* \\ -[\sigma b\lambda_3 + \omega_0^2(b + 1)] & \sigma(b\lambda_3 - \omega_0^2) & -\lambda_3\sigma u_1^* \\ \omega_0[\sigma(1 + \sigma) + \omega_0^2] & -\omega_0\sigma(1 + \sigma) & -\sigma\omega_0 u_1^* \end{bmatrix} \equiv \|q_{ij}\|. \quad (3.90)$$

Let us replace (3.89), (3.90), and (3.84) in (3.80) and compute the second-order partial derivatives of $F(x)$. It results in:

$$\begin{aligned} F_{x_1 x_1}^i &= 2(q_{i3}t_{21} - q_{i2}t_{31}) \\ F_{x_2 x_2}^i &= 0 \\ F_{x_3 x_3}^i &= 2(q_{i3}t_{23} - q_{i2}t_{33}) \\ F_{x_1 x_2}^i &= F_{x_2 x_3}^i = q_{i3}t_{22} - q_{i2}t_{32} \\ F_{x_1 x_3}^i &= q_{i3}(t_{21} + t_{23}) - q_{i2}(t_{31} + t_{33}), \quad i = 1, 2, 3 \\ F_{x_k x_l x_j}^i &= 0, \quad i, j, k, l = 1, 2, 3 \end{aligned} \quad (3.91)$$

TABLE 3.1Values for a_1 , $\alpha'(r_c)$, and δ

b, σ	a_1	$\alpha'(r_c)$	$\delta = -a_1/\alpha'(r_c)$
$b = 8/3, \sigma = 10$	3.86687×10^{-3}	3.02225×10^{-2}	-1.27947×10^{-1}
$b = 10, \sigma = 20$	9.47803×10^{-4}	2.37550×10^{-2}	-3.98990×10^{-2}
$b = 10, \sigma = 40$	1.01724×10^{-3}	3.88528×10^{-2}	-2.61818×10^{-2}
$b = 20, \sigma = 40$	4.63320×10^{-4}	2.64859×10^{-2}	-1.74931×10^{-2}
$b = 30, \sigma = 40$	2.40617×10^{-4}	8.45106×10^{-3}	-2.84718×10^{-2}

Substituting into (3.81) we obtain the Liapunov number a_1 . We have Table 3.1.

In this table $\alpha'(r_c)$ is given by (3.88). Notice that the values for δ correspond to those of μ_2 in Table 3.4 from [5]. For these values of b and σ , the bifurcation is subcritical and its direction is on the side $r < r_c$.

3.3.3 The Bifurcation in the Rössler System

In this section, we introduce the bifurcation in the Rössler system as an example of bifurcations of codimension 2. In general, our discussion follows Gaspard [3], but we will stop at the second order for the sake of simplicity.

The Rössler system reads:

$$\dot{u}_1 = -u_2 - u_3, \quad \dot{u}_2 = u_1 + au_2, \quad \dot{u}_3 = bu_1 - cu_3 + u_1u_3 \quad (3.92)$$

The system has two fixed points:

1. First fixed point:

$$O: \quad u_1 = u_2 = u_3 = 0$$

with the eigenvalues of the Jacobian matrix given by:

$$\lambda^3 + (c - a)\lambda^2 + (1 + b - ac)\lambda + (c - ab) = 0 \quad (3.93)$$

2. Second fixed point:

$$P: \quad u_1 = c - ab, \quad u_2 = b - \frac{c}{a}, \quad u_3 = \frac{c}{a} - b \quad (3.94)$$

with the eigenvalues of the Jacobian matrix given by:

$$\lambda^3 + a(b - 1)\lambda^2 + \left(1 + \frac{c}{a} - a^2b\right)\lambda - (c - ab) = 0 \quad (3.95)$$

The two fixed points coalesce on the surface $c = ab$ of the parameter space. The common eigenvalues are hence given by:

$$(\pm i\omega, 0) \quad \text{with } \omega = (2 - a^2)^{1/2} \text{ if } b = 1, \quad -\sqrt{2} \leq a \leq \sqrt{2} \quad (3.96)$$

The set of bifurcation points of this type is the segment

$$b = 1, \quad c = a, \quad -\sqrt{2} \leq a \leq \sqrt{2} \quad (3.97)$$

in the parameter space $\mathbb{R}^3$. It is hence a bifurcation of codimension 2. Let us seek the unfolding around the bifurcation point $b = 1, c = ab = a$; for this purpose, we set:

$$b = 1 + \varepsilon_1, \quad c = a + \varepsilon_2 \quad (3.98)$$

3.3.3.1 Reduction to the Normal Form

When $\varepsilon_1 = \varepsilon_2 = 0$, the system (3.92) reads:

$$\dot{u} = Au + \hat{F}(u) \quad (3.99)$$

with

$$A = \begin{bmatrix} 0 & -1 & -1 \\ 1 & a & 0 \\ 1 & 0 & -a \end{bmatrix}$$

and $\hat{F}(u) = (0, 0, u_1 u_3)^T$. The first step in the reduction procedure is to write (3.99) in the form (3.79) and (3.80). The eigenvalues of the matrix A are given by (3.96). With respect to the eigenvector basis,

$$v_1 = (2, -a - i\omega, a - i\omega), \quad v_2 = \bar{v}_1, \quad v_3 = (a, -1, 1)$$

the new variables, $\eta = (\eta_1, \eta_2, \eta_3)$, are related to the old variables through:

$$u = T\eta, \quad \eta = T^{-1}u$$

where T is the transformation matrix $[\Re v_1, -\Im v_1, v_3]$:

$$T = \begin{bmatrix} 2 & 0 & a \\ -a & \omega & -1 \\ a & \omega & 1 \end{bmatrix}, \quad T^{-1} = \frac{1}{\omega^2} \begin{bmatrix} 1 & \frac{a}{2} & -\frac{a}{2} \\ 0 & \frac{\omega}{2} & \frac{\omega}{2} \\ -a & -1 & 1 \end{bmatrix}$$

provided that $\omega^2 + a^2 = 2$. Equation (3.99) reads:

$$\dot{\eta} = A'\eta + F(\eta)$$

with

$$A' = T^{-1}AT = \begin{bmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad F(\eta) = T^{-1}\hat{F}(T\eta)$$

or, setting $\chi = \eta_1 + i\eta_2$

$$\dot{\chi} = i\omega\chi - \frac{a - i\omega}{2\omega^2}\Phi(\chi, \bar{\chi}, \eta_3), \quad \dot{\eta}_3 = \frac{1}{\omega^2}\Phi(\chi, \bar{\chi}, \eta_3) \quad (3.100)$$

with

$$\Phi(\chi, \bar{\chi}, \eta_3) = (\chi + \bar{\chi} + a\eta_3) \left[\frac{1}{2}(a - i\omega)\chi + \frac{1}{2}(a + i\omega)\bar{\chi} + \eta_3 \right]$$

To eliminate the nonresonant terms of order 2, we perform a quadratic transformation $\eta = \zeta + h^{(2)}\zeta\zeta$ which can be written in the form:

$$\begin{aligned} \chi &= \psi + \alpha_{11}\psi^2 + \alpha_{12}\psi\bar{\psi} + \alpha_{22}\bar{\psi}^2 + \alpha_{13}\psi\zeta_3 + \alpha_{23}\bar{\psi}\zeta_3 + \alpha_{33}\zeta_3^2 \\ \eta_3 &= \zeta_3 + \beta_{11}\psi^2 + \beta_{12}\psi\bar{\psi} + \beta_{22}\bar{\psi}^2 + \beta_{13}\psi\zeta_3 + \beta_{23}\bar{\psi}\zeta_3 + \beta_{33}\zeta_3^2 \end{aligned}$$

with $\psi = \zeta_1 + i\zeta_2$. We can now write (3.100) in the form:

$$(1 + A_1)\dot{\psi} = -A_2\dot{\bar{\psi}} - A_3\dot{\zeta}_3 + f(\psi, \bar{\psi}, \zeta_3) \quad (3.101)$$

$$(1 + B_1)\dot{\zeta}_3 = -B_2\dot{\psi} - B_3\dot{\bar{\psi}} + g(\psi, \bar{\psi}, \zeta_3) \quad (3.102)$$

where we have set $A_1 = 2\alpha_{11}\psi + \alpha_{12}\bar{\psi} + \alpha_{13}\zeta_3$, $A_2 = \alpha_{12}\psi + 2\alpha_{22}\bar{\psi} + \alpha_{23}\zeta_3$, $A_3 = \alpha_{13}\psi + \alpha_{23}\bar{\psi} + 2\alpha_{33}\zeta_3$, $B_1 = \beta_{13}\psi + \beta_{23}\bar{\psi} + 2\beta_{33}\zeta_3$, $B_2 = 2\beta_{11}\psi + \beta_{12}\bar{\psi} + \beta_{13}\zeta_3$, $B_3 = \beta_{12}\psi + 2\beta_{22}\bar{\psi} + \beta_{23}\zeta_3$, and $f(\psi, \bar{\psi}, \zeta_3)$, $g(\psi, \bar{\psi}, \zeta_3)$ as the right-hand sides of (3.100). At the order considered, we can replace: in the first relation, $\dot{\bar{\psi}}$ and $\dot{\zeta}_3$, respectively, with $-i\omega\bar{\psi}$ and 0; and in the second relation, $\dot{\psi}$ and $\dot{\bar{\psi}}$ with $i\omega\psi$ and $-i\omega\bar{\psi}$. Besides, to isolate $\dot{\psi}$ and $\dot{\zeta}_3$, we multiply (3.101) by $(1 - A_1)$ and (3.102) by $(1 - B_1)$ (see Example 3). After eliminating the nonresonant terms:

$$\begin{aligned} \alpha_{11} &= \frac{i(a - i\omega)^2}{4\omega^3}, \quad \alpha_{22} = -\frac{i}{6\omega^3}, \quad \alpha_{23} = -\frac{\omega + 2ia}{4\omega^3}, \\ \alpha_{12} &= \alpha_{33} = -\frac{a(\omega + ia)}{2\omega^3} = 2a\beta_{11} = 2a\bar{\beta}_{22}, \\ \beta_{13} &= -\frac{a\omega + i(2 + a^2)}{2\omega^3} = \bar{\beta}_{23}, \quad \alpha_{13}, \beta_{12}, \beta_{33} \text{ arbitrary} \end{aligned}$$

we obtain the following simplified system:

$$\begin{aligned}\dot{\psi} &= i\omega\psi - \frac{a}{\omega^2} \left[\frac{a^2}{2} - \frac{i\omega(1+a^2)}{2a} \right] \psi \zeta_3 \\ \dot{\zeta}_3 &= \frac{a}{\omega^2} \left(\zeta_3^2 + \psi \bar{\psi} \right), \quad \text{with } \psi = \zeta_1 + i\zeta_2\end{aligned}$$

where the resonant terms are independent of α_{ij} , β_{ij} . Eventually, by an appropriate change of scales

$$\zeta_1 = -\frac{\omega^2}{a}x, \quad \zeta_2 = -\frac{\omega^2}{a}y, \quad \zeta_3 = -\frac{\omega^2}{a}z$$

we obtain the Guckenheimer–Holmes normal form [4]:

$$\dot{q} = i\omega q + (\alpha + i\beta)zq, \quad \dot{z} = -z^2 - |q|^2, \quad q = x + iy \quad (3.103)$$

with

$$\alpha = \frac{a^2}{2}, \quad \beta = -\frac{\omega(a^2 + 1)}{2a} \quad (3.104)$$

The unfolding of (3.103) is, according to [4],

$$\dot{q} = (\mu_1 + i\omega)q + (\alpha + i\beta)zq, \quad \dot{z} = \mu_2 - z^2 - |q|^2 \quad (3.105)$$

To compute μ_1 , μ_2 as a function of ε_1 , ε_2 in (3.98), notice that the fixed points of (3.105) are:

$$P_{\pm} : \quad |q| = 0, \quad z = \pm\sqrt{\mu_2} \quad \text{if } \mu_2 > 0$$

It suffices thus to identify the eigenvalues at $P_{\pm}$ (computed as a function of μ_1 , μ_2) with the eigenvalues at O and P (computed as a function of ε_1 , ε_2).

In the coordinate system $(q, \bar{q}, z)$, the Jacobian matrix at the fixed points $P_{\pm}$ is diagonal:

$$J(0, 0, \pm\sqrt{\mu_2}) = \begin{bmatrix} \mu_1 + i\omega \pm (\alpha + i\beta)\sqrt{\mu_2} & 0 & 0 \\ 0 & \mu_1 - i\omega \pm (\alpha - i\beta)\sqrt{\mu_2} & 0 \\ 0 & 0 & \mp 2\sqrt{\mu_2} \end{bmatrix}$$

We may now compute the eigenvalues λ_1 , λ_2 , λ_3 for P_{+} :

$$\lambda_1 = \mu_1 + i\omega + (\alpha + i\beta)\sqrt{\mu_2}, \quad \lambda_2 = \bar{\lambda}_1, \quad \lambda_3 = -2\sqrt{\mu_2}$$

so that:

$$\lambda_1 + \lambda_2 + \lambda_3 = 2\mu_1 + 2(\alpha - 1)\sqrt{\mu_2}$$

Since we identify P_+ with O , $\lambda_1, \lambda_2, \lambda_3$ are also roots of (3.93), and thus:

$$2\mu_1 + 2(\alpha - 1)\sqrt{\mu_2} = -(c - a)$$

using the relations between roots and coefficients of a polynomial. We encounter an analogous relation for P_- :

$$2\mu_1 - 2(\alpha - 1)\sqrt{\mu_2} = -a(b - 1)$$

It follows that:

$$\mu_1 = -\frac{1}{4}(a\varepsilon_1 + \varepsilon_2), \quad \mu_2 = \frac{(a\varepsilon_1 - \varepsilon_2)^2}{4(a^2 - 2)^2}$$

where we have taken into account (3.98) and (3.104).

3.3.3.2 Bifurcation Diagram

Let us write (3.105) in cylindrical coordinates ($q = r e^{i\phi}, z$):

$$\dot{r} = (\mu_1 + \alpha z)r \tag{3.106a}$$

$$\dot{z} = \mu_2 - z^2 - r^2 \tag{3.106b}$$

$$\dot{\phi} = \omega + \beta z \tag{3.106c}$$

Equation (3.106c) can be decoupled from the others. The bifurcation diagram for the system (3.106a) and (3.106b) is displayed in Figure 3.4. Note

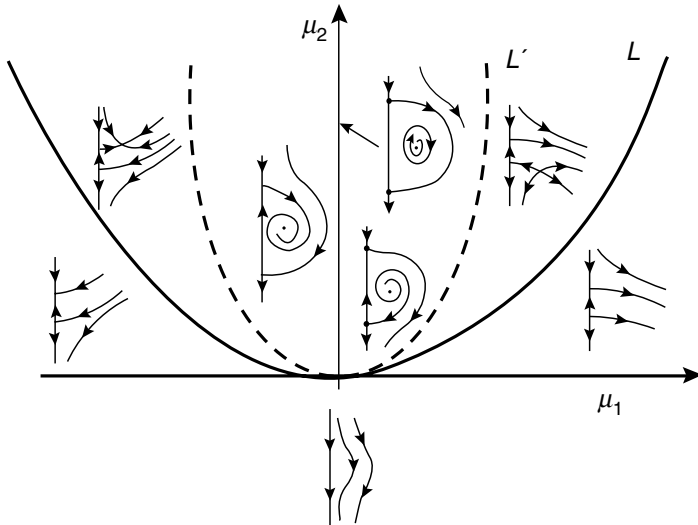
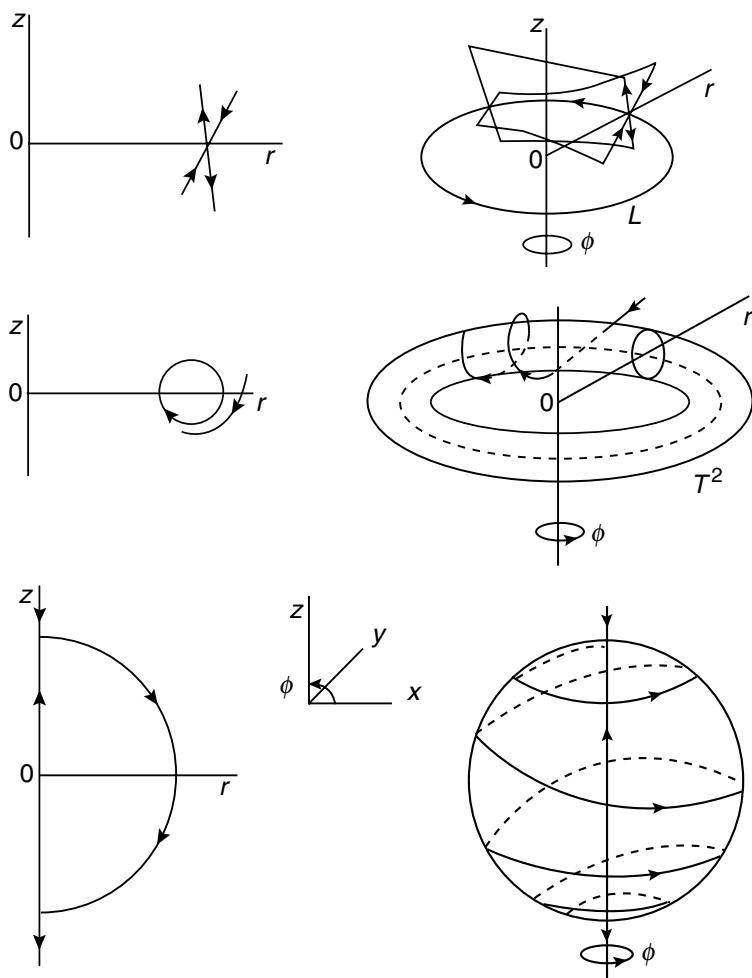


FIGURE 3.4

Bifurcation diagram for the Rössler model.

**FIGURE 3.5**

Correspondence between (r, z) and (r, ϕ, z) .

that a fixed point on the (r, z) plane corresponds to a periodic orbit, a limit cycle corresponds to a torus in three-dimensional space, and so forth (see Figure 3.5). If $\mu_2 > 0$, the system (3.106a) and (3.106b) has two fixed points:

$$r = 0, \quad z = \pm\sqrt{\mu_2}$$

and, inside the parabola L of equation,

$$\mu_2 = \frac{\mu_1^2}{\alpha^2}$$

a third fixed point:

$$r = \left(\mu_2 - \frac{\mu_1^2}{\alpha^2} \right)^{1/2}, \quad z = -\frac{\mu_1}{\alpha} \quad (3.107)$$

In cylindrical coordinates (r, ϕ, z) , the latter corresponds to a limit cycle.

The stability of the fixed points is determined by the eigenvalues of the Jacobian matrix

$$J(r, z) = \begin{pmatrix} \mu_1 + \alpha z & \alpha r \\ -2r & -2z \end{pmatrix} \quad (3.108)$$

at these points. For the fixed points $(r, z) = (0, \pm\sqrt{\mu_2})$, the matrix (3.108) is diagonal:

$$J(0, \pm\sqrt{\mu_2}) = \begin{pmatrix} \mu_1 \pm \alpha\sqrt{\mu_2} & 0 \\ 0 & \mp 2\sqrt{\mu_2} \end{pmatrix}$$

The classification of these fixed points is summarized in Table 3.2.

The matrix (3.108) for the fixed point (3.107) reads:

$$\begin{pmatrix} 0 & \sqrt{\alpha^2 \mu_2 - \mu_1^2} \\ -\frac{2}{\alpha} \sqrt{\alpha^2 \mu_2 - \mu_1^2} & 2\frac{\mu_1}{\alpha} \end{pmatrix}$$

This matrix has eigenvalues:

$$\lambda_{1,2} = \frac{\mu_1 \pm \sqrt{(1 + 2\alpha)\mu_1^2 - 2\alpha^3\mu_2}}{\alpha}$$

They are complex conjugates inside the parabola L' of equation:

$$(1 + 2\alpha)\mu_1^2 - 2\alpha^3\mu_2 = 0$$

TABLE 3.2

Classification of fixed points

(r, z)	$(0, +\sqrt{\mu_2})$	$(0, -\sqrt{\mu_2})$
$\mu_1 > \alpha\sqrt{\mu_2}$	Saddle	Source
$\alpha\sqrt{\mu_2} > \mu_1 > -\alpha\sqrt{\mu_2}$	Saddle	Saddle
$-\alpha\sqrt{\mu_2} > \mu_1$	Sink	Saddle

and we can easily remark that, along the axis $\mu_1 = 0, \mu_2 > 0$, there is a Hopf bifurcation at the fixed point (3.107) (see Figure 3.4). In fact, for $\mu_1 = 0$, the system (3.106a) and (3.106b) is integrable, with solution curves

$$\frac{\alpha r^{2/\alpha}}{2} \left(\mu_2 - \frac{r^2}{1 + \alpha} - z^2 \right) = C$$

References

1. R.I. Bogdanov, *Funct. Anal. Appl.*, 9, 144, 1975.
2. H. Dang-Vu and C. Delcarte, *Bifurcations et Chaos*, Ellipses, Paris, 2000.
3. P. Gaspard, *Physica*, 62D, 94, 1993.
4. J.A. Guckenheimer and P.J. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.
5. B.D. Hassard, N.D. Kazarinoff, and Y.H. Wan, *Theory and Applications of Hopf Bifurcation*, Cambridge University Press, Cambridge, 1981.
6. P. Manneville, *Systèmes Dynamiques et Chaos*, École Polytechnique, Palaiseau, 1999.
7. J.E. Marsden and M. McCracken, *The Hopf Bifurcation and Its Applications*, Springer-Verlag, Berlin, 1976.
8. F. Pham, *Géométrie et Calcul Différentiel sur les Variétés*, Dunod, Paris, 1999.
9. F. Takens, *Publ. Math. IHES*, 43, 47, 1974.
10. S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, Berlin, 1990.

4

Feedback Equivalence of Nonlinear Control Systems: A Survey on Formal Approach

W. Respondek and I. A. Tall

CONTENTS

4.1	Introduction	138
4.2	Equivalence of Dynamical Systems: Poincaré Theorem	147
4.3	Normal Forms for Single-Input Systems with Controllable Linearization	151
4.3.1	Introduction	151
4.3.2	Notations	152
4.3.3	Normal Form and m -Invariants	153
4.3.4	Normal Form for Non-affine Systems	162
4.4	Canonical Form for Single-Input Systems with Controllable Linearization	164
4.5	Dual Normal Form and Dual m -Invariants	172
4.6	Dual Canonical Form	176
4.7	Normal Forms for Single-Input Systems with Uncontrollable Linearization	178
4.7.1	Introduction	178
4.7.2	Taylor Series Expansions	179
4.7.3	Linear Part and Resonances	181
4.7.4	Notations and Definitions	182
4.7.5	Weighted Homogeneous Systems	185
4.7.6	Weighted Homogeneous Invariants	190
4.7.7	Explicit Normalizing Transformations	192
4.7.8	Weighted Normal Form for Single-Input Systems with Uncontrollable Linearization	193
4.7.9	Example	195
4.8	Normal Forms for Multi-Input Nonlinear Control Systems	198
4.8.1	Non-affine Normal Forms	200
4.8.2	Affine Normal Forms	202
4.8.3	Examples	203

4.9	Feedback Linearization	210
4.10	Normal Forms for Discrete Time Control Systems	215
4.10.1	Example: Bressan and Rampazzo Pendulum	219
4.11	Symmetries of Control Systems	221
4.11.1	Symmetries	222
4.11.2	Symmetries of Single-Input Nonlinearizable Systems	223
4.11.3	Symmetries of the Canonical Form	225
4.11.4	Formal Symmetries	225
4.11.5	Symmetries of Feedback Linearizable Systems	228
4.12	Feedforward and Strict Feedforward Forms	232
4.12.1	Introduction and Notations	232
4.12.2	Feedforward and Strict Feedforward Normal Forms	235
4.12.3	Feedforward and Strict Feedforward Form: First Nonlinearizable Term	237
4.12.4	Feedforward and Strict Feedforward Forms: The General Step	239
4.12.5	Feedforward and Strict Feedforward Systems on $\mathbb{R}^4$	242
4.12.5.1	Feedforward Case	242
4.12.5.2	Strict Feedforward Case	244
4.12.6	Geometric Characterization of Feedforward and Strict Feedforward Systems	246
4.12.7	Symmetries and Strict Feedforward Form	247
4.12.8	Strict Feedforward Form: Affine Versus General	251
4.12.9	Strict Feedforward Systems on the Plane	252
4.13	Analytic Normal Forms: A Class of Strict Feedforward Systems	253
4.14	Conclusions	256
	References	257

4.1 Introduction

In this chapter, we will deal with nonlinear control systems of the form:

$$\Pi : \dot{x} = F(x, u)$$

where $x \in X$ is an open subset of $\mathbb{R}^n$ and $u \in U \subset \mathbb{R}^m$, and $F(x, u)$ is a family of vector fields, C^∞ -smooth, with respect to (x, u) . The variables $x = (x_1, \dots, x_n)^T$ represent the state of the system and the variables $u = (u_1, \dots, u_m)^T$ represent the control (i.e., an external influence on the system). Π can be understood as underdetermined system of ordinary differential equations: n equations for $n + m$ variables.

We are interested in equivalence problems for the system Π . Consider another system of the same form:

$$\tilde{\Pi} : \dot{\tilde{x}} = \tilde{F}(\tilde{x}, \tilde{u})$$

where $\tilde{x} \in \tilde{X}$ is an open subset of $\mathbb{R}^n$ and $\tilde{u} \in \tilde{U} \subset \mathbb{R}^m$. A natural equivalence can be defined as follows. Assume that $U = \tilde{U}$. We say that Π and $\tilde{\Pi}$ are *state-space equivalent* (*S-equivalent*), if there exist a diffeomorphism

$$\tilde{x} = \phi(x)$$

$$\tilde{u} = u$$

transforming solutions into solutions. More precisely, if $(x(t), u(t))$ is a solution of Π , then $(\phi(x(t)), u(t))$ is a solution of $\tilde{\Pi}$, which is equivalent to

$$D\phi(x) \cdot F(x, u) = \tilde{F}(\phi(x), u)$$

for any $u \in U$, where $D\phi(x)$ denotes the derivative of ϕ at x . This means that the S-equivalent establishes a diffeomorphic correspondence of the right-hand sides of differential equations corresponding to the constant controls, which can be expressed as

$$(\phi_* F)(\tilde{x}, u) = \tilde{F}(\tilde{x}, u), \quad u \in U$$

where, for any vector field f and any diffeomorphism $\tilde{x} = \phi(x)$, we denote $(\phi_* f)(\tilde{x}) = D\phi(\phi^{-1}(\tilde{x})) \cdot f(\phi^{-1}(\tilde{x}))$.

S-equivalence is well understood. It establishes a one-to-one smooth correspondence between the trajectories of equivalent systems (corresponding to the same measurable, not necessarily constant, controls). For accessible systems [38, 66], the set of complete invariants for the local S-equivalence is formed by all iterative Lie-brackets evaluated at a nominal point (in the analytical case) or in its neighborhood (in the smooth case) [40].

Since the system Π has state *and* control variables, another natural transformation is to apply to Π a diffeomorphism $\Upsilon = (\phi, \psi)^T$ of $X \times U$ onto $\tilde{X} \times \tilde{U}$ that changes both x and u , that is:

$$\tilde{x} = \phi(x, u)$$

$$\tilde{u} = \psi(x, u)$$

and transforms the solutions of Π into those of $\tilde{\Pi}$. Taking a C^1 -solution $(x(t), u(t))$ of Π and using the fact that its image $(\phi(x(t), u(t)),$

$\psi(x(t), u(t))$ is assumed to be a solution of $\tilde{\Pi}$, we conclude that $(\partial\phi/\partial x)F(x, u) + (\partial\phi/\partial u)\dot{u} = \tilde{F}(\phi(x, u), \psi(x, u))$. Now, it is easy to see that, since F and $\tilde{F}$ do not depend on, respectively, $\dot{u}$ and $\dot{\tilde{u}}$, the map ϕ cannot depend on u . This implies that any Υ preserving the system solutions must actually be a triangular diffeomorphism

$$\Upsilon : \begin{aligned} \tilde{x} &= \phi(x) \\ \tilde{u} &= \psi(x, u) \end{aligned}$$

satisfying

$$D\phi(x) \cdot F(x, u) = \tilde{F}(\phi(x), \psi(x, u)),$$

which is called a *feedback transformation*. Systems Π and $\tilde{\Pi}$, equivalent through Υ , are called *feedback equivalent* (*F-equivalent*). The states x and $\tilde{x}$ of two feedback-equivalent systems are thus related by a diffeomorphism ϕ between the corresponding state-spaces X and $\tilde{X}$, whereas the controls u and $\tilde{u}$ are related by a diffeomorphism ψ between U and $\tilde{U}$ which depends on the state x . We will call Π and $\tilde{\Pi}$ locally feedback equivalent at (x_0, u_0) and $(\tilde{x}_0, \tilde{u}_0)$, respectively, if (ϕ, ψ) is a local diffeomorphism satisfying $(\phi, \psi)(x_0, u_0) = (\tilde{x}_0, \tilde{u}_0)$. The feedback equivalence and its local counterpart are the main topics of this chapter.

Note that the diffeomorphism ϕ establishes a one-to-one correspondence of x -trajectories of two feedback-equivalent systems although equivalent trajectories are differently parameterized by controls. Indeed, a trajectory $x(t)$ of the first system corresponding to a control $u(t)$ is mapped into the curve $\phi(x(t))$, which is the trajectory of $\tilde{\Pi}$ corresponding to $\tilde{u}(t) = \psi(x(t), u(t))$. On the basis of this observation, one can define a weaker notion of equivalence of Π and $\tilde{\Pi}$ asking that there exists a one-to-one correspondence between trajectories (corresponding to, e.g., C^∞ -controls) and omitting the assumption that the correspondence is given by a diffeomorphism. This leads to the important notion of dynamic feedback equivalence [19, 20, 41, 68], which, however, we will not discuss in this chapter.

The main subject of this chapter is feedback equivalence, which has been extensively studied during the last 20 years. Although natural, this problem is very involved (mainly because of the functional parameters appearing in the classification that will be explained briefly below). Many existing results are devoted to systems that are affine with respect to controls, that is, are of the form

$$\Sigma : \dot{x} = f(x) + \sum_{i=1}^m g_i(x)u_i = f(x) + g(x)u$$

where $x \in X$, f and g_i are C^∞ -smooth control vector fields on X , $u = (u_1, \dots, u_m)^T \in U = \mathbb{R}^m$ and $g = (g_1, \dots, g_m)$. When studying the feedback

equivalence of control-affine systems, we will apply feedback transformations that are affine with respect to controls:

$$\begin{aligned}\Gamma : \quad & \tilde{x} = \phi(x) \\ & u = \alpha(x) + \beta(x)\tilde{u}\end{aligned}$$

where $u = \psi^{-1}(x, \tilde{u}) = \alpha(x) + \beta(x)\tilde{u}$, with α and β being C^∞ -smooth functions with values in $\mathbb{R}^m$ and $Gl(m, \mathbb{R})$, respectively. Consider another control-affine system

$$\tilde{\Sigma} : \dot{\tilde{x}} = \tilde{f}(\tilde{x}) + \sum_{i=1}^m \tilde{g}_i(\tilde{x})\tilde{u}_i = \tilde{f}(\tilde{x}) + \tilde{g}(\tilde{x})\tilde{u}$$

where $\tilde{u} = (\tilde{u}_1, \dots, \tilde{u}_m)^T \in \tilde{U} = \mathbb{R}^m$ and $\tilde{g} = (\tilde{g}_1, \dots, \tilde{g}_m)$.

The general definition implies that the control-affine systems Σ and $\tilde{\Sigma}$ are feedback equivalent if and only if

$$\phi_*(f + g\alpha) = \tilde{f} \quad \text{and} \quad \phi_*(g\beta) = \tilde{g}$$

which we will write as

$$\Gamma_*(\Sigma) = \tilde{\Sigma}$$

We will say that the control-affine systems Σ and $\tilde{\Sigma}$ are locally feedback equivalent at x_0 and $\tilde{x}_0$, respectively, if ϕ is a local diffeomorphism satisfying $\phi(x_0) = \tilde{x}_0$ and α and β are defined locally around x_0 . Note that local feedback equivalence is local in the state-space X but global in the control space $U = \mathbb{R}^m$.

Given two control-affine systems Σ and $\tilde{\Sigma}$, the problem of their (local) feedback equivalence amounts to solving the system of first-order partial differential equations

$$\begin{aligned}\frac{\partial \phi}{\partial x}(x)(f(x) + g(x)\alpha(x)) &= \tilde{f}(\phi(x)) \\ \frac{\partial \phi}{\partial x}(x)(g(x)\beta(x)) &= \tilde{g}(\phi(x))\end{aligned}\tag{CDE}$$

Feedback equivalence of general systems Π under Υ and of control-affine systems Σ under Γ are very closely related. Consider a general nonlinear control system

$$\Pi : \dot{x} = F(x, u)$$

where $x \in X$, an open subset of $\mathbb{R}^n$ and $u \in U$, an open subset of $\mathbb{R}^m$. Together with Π , its extension (preintegration)

$$\Sigma^e : \dot{x}^e = f^e(x^e) + g^e(x^e)u^e$$

where $x^e = (x, u) \in X^e = X \times U$, $u^e \in U^e = \mathbb{R}^m$, and the dynamics are given by

$$\dot{x} = F(x, u)$$

$$\dot{u} = u^e$$

that is, $f^e(x^e) = (F(x, u), 0)^T$ and $g^e(x^e) = (0, Id)^T$. Notice that Σ^e is a control-affine system controlled by the derivatives $\dot{u}_i = u_i^e$ of the original controls u_i , for $1 \leq i \leq m$.

PROPOSITION 1

Two control systems Π and $\tilde{\Pi}$ are equivalent (resp. locally equivalent at (x_0, u_0) and $(\tilde{x}_0, \tilde{u}_0)$) under a general feedback transformation Υ if and only if their respective extensions Σ^e and $\tilde{\Sigma}^e$ are equivalent (resp. locally equivalent at $x_0^e = (x_0, u_0)$ and $\tilde{x}_0^e = (\tilde{x}_0, \tilde{u}_0)$) under an affine feedback Γ .

As a consequence, many problems concerning feedback equivalence are studied and solved for control-affine systems and their extension to the general case can be done by an appropriate application of Proposition 1.

To geometrize the problem of feedback equivalence, we associate its *field of admissible velocities* to the system Π

$$\mathcal{F}(x) = \{F(x, u) : u \in U\} \subset T_x X$$

The field of admissible velocities of the control-affine system Σ is the following field of affine subspaces (equivalently, an affine distribution):

$$\mathcal{A}(x) = \left\{ f(x) + \sum_{i=1}^m g_i(x)u_i : u_i \in \mathbb{R} \right\} = f(x) + \mathcal{G}(x) \subset T_x X$$

where $\mathcal{G}$ denotes the distribution spanned by the vector fields $g_1, \dots, g_m$. Now it is easy to see that if two control affine-systems Σ and $\tilde{\Sigma}$ are feedback equivalent, then the corresponding affine distributions are equivalent, that is:

$$\phi_* \mathcal{A} = \tilde{\mathcal{A}}$$

Moreover, the converse holds if the distributions $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are of constant rank m . Analogous implications (the converse under the constant rank

assumption) are true for local feedback equivalence. Note that attaching the field of admissible velocities to an affine system results in eliminating controls from the description: what remains is a geometric object, which is the affine distribution $\mathcal{A}$ while the choice of controls (equivalently, the choice of sections of $\mathcal{A}$) becomes irrelevant.

Example 1

To illustrate the notion of feedback equivalence, recall the first (historically) studied feedback classification problem, which is that for linear control systems of the form

$$\Lambda : \dot{x} = Ax + Bu = Ax + \sum_{i=1}^m u_i b_i$$

where $x \in \mathbb{R}^n$, Ax and $b_1, \dots, b_m$ are, respectively, linear and constant vector fields on $\mathbb{R}^n$, and $u = (u_1, \dots, u_m)^T \in \mathbb{R}^m$. To preserve the linear form of the system, we apply to it the linear feedback transformation.

$$\tilde{x} = Tx$$

$$u = Kx + L\tilde{u}$$

where T , K , and L are matrices of appropriate sizes (T and L being invertible). The system Λ is transformed into

$$\tilde{\Lambda} : \dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}\tilde{u} = T(A + BK)T^{-1}\tilde{x} + TBL\tilde{u}$$

It is a classical result of the linear control theory [49] that any linear controllable system is feedback equivalent to the following system (called Brunovský canonical form):

$$\begin{aligned} \dot{\tilde{x}}_{i,j} &= \tilde{x}_{i,j+1} & 1 \leq j \leq \rho_i - 1 \\ \dot{\tilde{x}}_{i,\rho_i} &= \tilde{u}_i & 1 \leq i \leq m \end{aligned}$$

where $\tilde{x} = (\tilde{x}_{1,1}, \tilde{x}_{1,2}, \dots, \tilde{x}_{1,\rho_1}, \dots, \tilde{x}_{m,\rho_m})^T$.

The integers $\rho_1 \geq \dots \geq \rho_m$, $\sum_{i=1}^m \rho_i = n$ (called controllability indices, Brunovský indices, or Kronecker indices) form a set of complete feedback invariants of the linear feedback linear group action on controllable systems and are defined as follows

$$\rho_i = \{q_j \mid q_j \geq i\} \quad (4.1)$$

where $m_0 = 0$, $m_i = \text{rank}(B, \dots, A^{i-1}B)$ and $q_i = m_i - m_{i-1}$ for $1 \leq i \leq n$.

Note that for the linear control system Λ , the field of admissible velocities is given by the field of m -dimensional affine subspaces $\mathcal{A}(x) = Ax + \mathcal{B}$ of $\mathbb{R}^n$, where $\mathcal{B}$ is the image of $\mathbb{R}^m$ under the linear map $B : \mathbb{R}^m \rightarrow \mathbb{R}^n$. The Brunovský canonical form thus gives a canonical form for the field $\mathcal{A}$ under linear invertible transformations $\tilde{x} = Tx$.

Observe that the dimension of the space of linear systems (of pairs (A, B)) is $n^2 + nm$, and the dimension of the group of linear feedback (of the triples (T, K, L)) is $n^2 + nm + m^2$. We can thus expect open orbits to exist and, indeed, they do (those of systems with the maximal vector of d_i 's).

The picture gets completely different for nonlinear control systems under the action of nonlinear feedback. Although both are infinite dimensional, the group of (local) feedback transformations is much “smaller” than the space of all (local) control systems and, as a consequence, functional parameters must necessarily appear in the feedback classification. To observe this, note that the space of general systems Π is parameterized by n functions of $n + m$ variables (components of F) while the group of feedback transformations by m functions of $n + m$ variables (components of ψ) and n functions of n variables (components of ϕ). Thus, functional parameters are to be expected if $m < n$, that is, in all interesting cases. To make this argument precise, we will follow Ref. [40] and compute the dimension $d_\Pi(k)$ of the space of k -jets of the system Π and the dimension $d_\Upsilon(k)$ of the corresponding jet-space of the feedback group acting on the space of k -jets of the systems. To this end, recall that the dimension of the space of polynomials of n variables, of degree not greater than k , is equal to

$$\frac{(k+n)!}{k!n!} = \frac{(k+1)(k+2) \cdots (k+n)}{n!}$$

which is a polynomial of k of degree n starting with $k^n/n!$. The dynamics of the system are represented by n components of $F(x, u)$, each being a function of $n + m$ variables. The feedback group is represented by n components of the diffeomorphism $\phi(x)$, each being a function of n variables and m components of the map $\psi(x, u)$, each being a function of $n + m$ variables. Note that the $(k+1)$ -jet of a diffeomorphism acts on the k -jet of the system. Thus

$$d_\Pi(k) = n \frac{(k+n+m)!}{k!(n+m)!}, \quad d_\Upsilon(k) = n \frac{(k+1+n)!}{(k+1)!n!} + m \frac{(k+n+m)!}{k!(n+m)!}$$

The codimension of any orbit, of the feedback group action on the space of systems, is bounded from below by the difference $d_\Pi(k) - d_\Upsilon(k)$, which is a polynomial of k of degree $n + m$, whose coefficient multiplying k^{n+m} is

$$\frac{n-m}{(n+m)!}$$

This coefficient is positive when $m < n$, and thus the polynomial and the codimension of any orbit tend to infinity when k tends to infinity. As a consequence, functional moduli must appear in the feedback classification if $m < n$ (which exhausts all interesting cases).

Observe that a control-affine system is defined by $m + 1$ vector fields $f, g_1, \dots, g_m$ and the feedback group by the diffeomorphism ϕ and $m + m^2$ components of the pair (α, β) . Therefore, in the case of control-affine systems the corresponding dimensions are:

$$d_\Sigma = n(m+1) \frac{(k+n)!}{k!(n)!}, \quad d_\Gamma = n \frac{(k+1+n)!}{(k+1)!n!} + m(m+1) \frac{(k+n)!}{k!(n)!}$$

The codimension of any orbit in the k -jets space is bounded below by the difference $d_\Sigma - d_\Gamma$, which is a polynomial of k of degree n , whose coefficient multiplying k^n is

$$\frac{m(n-m-1)}{n!}$$

If $m < n - 1$, then this coefficient is positive and thus the polynomial and the codimension of any orbit under the feedback group action tend to infinity as k tends to infinity. As a consequence, functional moduli must appear in feedback classification of control-affine systems if $m < n - 1$. In the case $m = n - 1$ we can hope, however, for normal forms without functional parameters and, indeed, such normal forms have been obtained by Respondek and Zhitomirskii both for $m = 2, n = 3$ [75] and for the general case [95].

It is the existence of functional moduli which causes one of the main difficulties of the feedback equivalence problem. Four basic methods have been proposed to study various aspects of feedback equivalence. The first method, used for control-affine systems, is based on studying invariant properties of two geometric objects attached to the system: the distribution $\mathcal{G}$ and the affine distribution $\mathcal{A}$. Note that feedback equivalence of control-linear systems (i.e., control-affine system Σ with $f \equiv 0$) coincides with equivalence, under a diffeomorphism, of the corresponding distributions $\mathcal{G}$ and $\tilde{\mathcal{G}}$. Thus this approach is linked, in a natural way, with the classification and with singularities of vector fields and distributions, and their invariants. Using this method a large variety of feedback classification problems have been solved [9, 11, 14, 37, 40, 45, 46, 56, 69, 75, 95].

The second approach, proposed by Gardner [21], uses Cartan's method of equivalence [13]. To the control system Π , we can associate the Pfaffian system given by the differential forms $dx_i - F_i(x, u) dt$, for $1 \leq i \leq n$, on $X \times U \times \mathbb{R}$, and the feedback equivalence of Π and $\tilde{\Pi}$ is analyzed by studying the equivalence of the corresponding Pfaffian systems and their geometry [23–25, 62].

The third method, inspired by the hamiltonian formalism for optimal control problems, has been developed by Bonnard and Jakubczyk [8, 9, 42, 44] and has led to a very nice description of feedback invariants in terms of singular extremals. Another approach based also on the hamiltonian formalism for optimal control has been proposed by Agrachev [1, 2] and has led to a construction of a fundamental geometric invariant of feedback equivalence: the curvature of control systems.

Finally, a very fruitful approach was proposed by Kang and Krener [54] and then followed by Kang [50, 51]. Their idea, which is closely related to the classical Poincaré's technique for linearization of dynamical systems [3], is to analyze the system Π and the feedback transformation Υ (the system Σ and the transformation Γ , respectively, in the control-affine case) step by step and, as a consequence, to produce a simpler equivalent system $\tilde{\Pi}$ also step by step. It is this approach, and various classification results obtained using it, which form the subject of this chapter.

This chapter is organized as follows. We will present in Section 4.2 the classical Poincaré's approach to the problem of formal equivalence of dynamical systems. In Section 4.3, we will generalize, following Kang and Krener, the formal approach to nonlinear control systems. We will present a normal form for homogeneous systems, their invariants, explicit normalizing transformations and, finally, a normal form under a formal feedback. We will also extend the normal form to general non-affine systems. Then, in Section 4.4, we will propose a canonical form for nonlinear control systems. In the following two sections (Section 4.5 and Section 4.6) we will dualize results of preceding sections and present a dual normal form (together with dual invariants and explicit normalizing transformations) and a dual canonical form. Then, in Section 4.7, we will pass to systems with uncontrollable linearization; introduce weighted homogeneity; and present a normal form, invariants, explicit normalizing transformations, and a formal normal form. This section generalizes, on the one hand, results of systems with controllable linearization (presented in earlier sections) and, on the other hand, results on dynamical systems from Section 4.2. Section 4.8 will be devoted to multi-input normal forms (for space-related reasons we treat only the controllable case): it generalizes results on normal forms obtained in Section 4.3. A discrete time version of Section 4.3 will be given in Section 4.10. In Section 4.9, we compare well-known results devoted to feedback linearization with their counterpart obtained via the formal feedback. We will also discuss systems that are feedback equivalent to linear uncontrollable systems. Then the following two sections present applications of the formal approach to the classification of control systems. We discuss symmetries of control systems in Section 4.11 and show an enormous difference between the group of symmetries of feedback linearizable and nonlinearizable systems. In Section 4.12, we characterize, using the formal approach, systems that are feedback equivalent to feedforward

and strict feedforward forms. Finally, in Section 4.13, we present a class of analytic strict feedforward forms than can be transformed to a normal form via constructive analytic transformations.

Because of space limit, this chapter does not touch many important results. To mention just a few: we do not discuss analysis of bifurcations based on formal approach [52, 53, 59, 60], bifurcations of discrete time systems, or normal forms for observed dynamics. Each of those subjects requires its own survey, proving the efficiency of the formal approach.

4.2 Equivalence of Dynamical Systems: Poincaré Theorem

In this section, we will summarize very briefly Poincaré's approach to the problem of (formal) equivalence of dynamical systems. The goal of this section is three-fold. First, to make our survey complete and self-contained. Secondly, to show how the formal approach to the equivalence of dynamical systems generalizes to the formal approach to feedback equivalence of control systems. Thirdly, some of results on formal normal forms for dynamical systems and of formal linearization (Theorem 1 and Theorem 2 stated at the end of this section) will be used in Section 4.7 and Section 4.9 of the survey.

Consider the uncontrolled dynamical system

$$\dot{x} = f(x)$$

where $x \in X$, an open subset of $\mathbb{R}^n$ and f is a C^∞ -smooth vector field on X . A C^∞ -smooth diffeomorphism

$$\tilde{x} = \phi(x)$$

brings the considered dynamical system into

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}) = (\phi_* f)(\tilde{x})$$

where

$$(\phi_* f)(\tilde{x}) = \frac{\partial \phi}{\partial x}(\phi^{-1}(\tilde{x})) \cdot f(\phi^{-1}(\tilde{x}))$$

Now given two dynamical systems $\dot{x} = f(x)$ and $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$, the problem of establishing their equivalence is to find a diffeomorphism $\tilde{x} = \phi(x)$ satisfying

$$\frac{\partial \phi}{\partial x}(x) \cdot f(x) = \tilde{f}(\phi(x)) \quad (\text{DE})$$

which is a system of n first-order partial differential equations for the components of $\phi(x)$. Notice that in the most interesting cases of $f(x_0) = \tilde{f}(\tilde{x}_0) = 0$, this is a system of singular partial differential equations.

Consider the infinite Taylor series expansion of our dynamical system

$$\dot{x} = f(x) = Jx + \sum_{m=2}^{\infty} f^{[m]}(x)$$

around an equilibrium, which is assumed to be $x_0 = 0 \in \mathbb{R}^n$, where $f^{[m]}$ denotes a polynomial vector field, all of whose components are homogenous polynomials of degree m . Apply to it a formal change of coordinates given by an invertible formal transformation of the form

$$\tilde{x} = \phi(x) = x + \sum_{m=2}^{\infty} \phi^{[m]}(x)$$

which preserves $0 \in \mathbb{R}^n$ and starts with the identity, where all components of $\phi^{[m]}$ are homogeneous polynomials of degree m . To study the action of $\phi(x)$ on $f(x)$, we will see how its homogenous part of degree m acts on terms of degree m of f . To this end, apply to

$$\dot{x} = Jx + f^{[m]}(x)$$

the transformation

$$\tilde{x} = x + \phi^{[m]}(x)$$

where $m \geq 2$. We have, modulo terms of higher degree,

$$\begin{aligned} \dot{\tilde{x}} &= Jx + f^{[m]}(x) + \frac{\partial \phi^{[m]}}{\partial x}(x) (Jx + f^{[m]}(x)) \\ &= J\tilde{x} - J\phi^{[m]}(x) + f^{[m]}(x) + \frac{\partial \phi^{[m]}}{\partial x}(x) Jx \\ &= J\tilde{x} + f^{[m]}(x) + [Jx, \phi^{[m]}(x)] \\ &= J\tilde{x} + \tilde{f}^{[m]}(\tilde{x}) \end{aligned}$$

where $[v, w](x) = (\partial w / \partial x)(x)v(x) - (\partial v / \partial x)(x)w(x)$ is the Lie bracket of two vector fields v and w . Using the notation $\text{ad}_v w = [v, w]$, we obtain

$$\text{ad}_{Jx} \phi^{[m]}(x) = \tilde{f}^{[m]}(x) - f^{[m]}(x) \quad (\text{HE})$$

which we will call a *homological equation*.

Consider the action of ad_{Jx} on the space $P^{[m]}$ of polynomial vector fields, all of whose components are homogeneous polynomials of degree m . For a multi-index $k = (k_1, \dots, k_n)$, denote $x^k = x_1^{k_1} \dots x_n^{k_n}$.

LEMMA 1

Assume that J is diagonal, say $J = \text{diag}(\lambda_1, \dots, \lambda_n)$. Then ad_{Jx} is a diagonal operator on the space $P^{[m]}$ in the eigenbasis formed by the eigenvectors $x^k (\partial/\partial x_i)$, for all multi-indices k such that $k_1 + \dots + k_n = m$ and $1 \leq i \leq n$. The eigenvalues of ad_{Jx} depend linearly on the eigenvalues of J , more precisely, we have

$$\text{ad}_{Jx} \left(x^k \frac{\partial}{\partial x_i} \right) = (k, \lambda) \left(x^k \frac{\partial}{\partial x_i} \right)$$

where $\lambda = (\lambda_1, \dots, \lambda_n)$, and $(k, \lambda) = k_1 \lambda_1 + \dots + k_n \lambda_n$.

COROLLARY 1

The operator ad_{Jx} is invertible on the space $P^{[m]}$ if there does not hold any relation of the form

$$\sum_{s=1}^n k_s \lambda_s = \lambda_j$$

where k_s are nonnegative integers, $|k| = k_1 + \dots + k_n \geq 2$ and $1 \leq j \leq n$.

For any relation $\lambda_j = \sum_{s=1}^n k_s \lambda_s$, called *resonance*, we define

$$\mathcal{R}_j = \{k = (k_1, \dots, k_n) : \lambda_j = k_1 \lambda_1 + \dots + k_n \lambda_n, \quad k_i \in \mathbb{N} \cup \{0\}, \quad |k| \geq 2\},$$

which will be called the *resonant set* associated with λ_j .

THEOREM 1

Consider the differential equation

$$\dot{x} = f(x) = Jx + \sum_{m=2}^{\infty} f^{[m]}(x)$$

and assume that all eigenvalues are real and distinct, and that the spectrum of J is nonresonant:

1. For each $m \geq 2$ and any homogenous vector fields $f^{[m]}$ and $\tilde{f}^{[m]}$ of degree m , the homological equation (HE) is solvable within the class of $\mathbb{R}^n$ -valued homogeneous polynomials $\phi^{[m]}$ of degree m .

2. The differential equations $\dot{x} = f(x) = Jx + \sum_{m=2}^{\infty} f^{[m]}(x)$ and $\dot{\tilde{x}} = \tilde{f}(\tilde{x}) = J\tilde{x} + \sum_{m=2}^{\infty} \tilde{f}^{[m]}(\tilde{x})$ are equivalent through an invertible formal transformation of the form $\tilde{x} = x + \sum_{m=2}^{\infty} \phi^{[m]}(x)$.
3. The differential equation $\dot{x} = f(x) = Jx + \sum_{m=2}^{\infty} f^{[m]}(x)$ can be formally linearized, that is, can be brought to the form $\dot{\tilde{x}} = J\tilde{x}$ through an invertible formal transformation of the form $\tilde{x} = x + \sum_{m=2}^{\infty} \phi^{[m]}(x)$.

Item (1) is a direct consequence of Corollary 1. Item (2) follows by a successive application of (1) for $m = 2, 3$, and so on. Finally, (3) is an immediate consequence of (2), applied for $\tilde{f} = J\tilde{x}$.

If the spectrum of J is resonant, then using the ad_J operator we can get rid of all nonresonant terms, which leads to the following:

THEOREM 2

Consider the differential equation

$$\dot{x} = f(x) = Jx + \sum_{m=2}^{\infty} f^{[m]}(x)$$

Assume that J is diagonal, that is, $J = \text{diag}(\lambda_1, \dots, \lambda_n)$. There exists a formal invertible transformation of the form $\tilde{x} = x + \sum_{m=2}^{\infty} \phi^{[m]}(x)$ bringing $\dot{x} = f(x)$ into $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$ of the form

$$\tilde{f}_j(\tilde{x}) = \lambda_j \tilde{x}_j + \sum_{k \in \mathcal{R}_j} \gamma_j^k \tilde{x}_1^{k_1} \dots \tilde{x}_n^{k_n}$$

where $\gamma_j^k \in \mathbb{R}$ and the summation is taken over all resonances $k = (k_1, \dots, k_n)$ forming the resonant set $\mathcal{R}_j$ associated with λ_j .

If the eigenvalues of J are distinct but not necessarily real, then an analogous result holds (which will be stated it in Section 4.7).

Theorem 1 and Theorem 2 summarize Poincaré's approach in the formal category. The idea of this approach is very natural: in order to establish the equivalence of two dynamical systems, we replace the singular partial differential equation (DE) by an infinite sequence of homological equations (HE), which are simply linear equations with respect to the unknown components of the homogenous part $\phi^{[m]}$ of ϕ .

Much more delicate and difficult issues of constructing C^∞ -smooth or real analytical transformations that linearize the equation (in the non-resonant case) or annihilate all nonresonant terms (in the general case) are discussed very briefly in Section 4.9.

4.3 Normal Forms for Single-Input Systems with Controllable Linearization

4.3.1 Introduction

In this section, we will study nonlinear single-input control-affine systems of the form

$$\Sigma : \dot{\xi} = f(\xi) + g(\xi)u$$

where $\xi \in X$, an open subset of $\mathbb{R}^n$, $u \in \mathbb{R}$, and f and g are C^∞ -smooth vector fields on X .

Throughout this section we will study the system Σ around a point ξ_0 at which $f(\xi_0) = 0$ and $g(\xi_0) \neq 0$. Without loss of generality, we will assume that $\xi_0 = 0$. We will also assume throughout this section that the linear part (F, G) of the system is controllable, where $F = (\partial f / \partial \xi)(0)$ and $G = g(0)$.

The goal of this section is to obtain a normal form of Σ under the action of the feedback group consisting of feedback transformation of the form

$$\Gamma : \begin{aligned} x &= \phi(\xi) \\ u &= \alpha(\xi) + \beta(\xi)v \end{aligned}$$

Together with the system Σ and the feedback transformation Γ , we will consider their Taylor series expansions Σ^∞ and Γ^∞ , respectively, and we will study the action of Γ^∞ on Σ^∞ step-by-step, that is, the action of the homogeneous part Γ^m of Γ^∞ on the homogeneous part $\Sigma^{[m]}$ of Σ^∞ . In other words, we will generalize the approach that Poincaré has developed for dynamical systems (which we recalled in Section 4.2) to control systems. It was Kang and Krener [50, 51, 54] who proposed this approach in the context of control systems and who have obtained fundamental results. Their pioneering work has inspired the authors who have obtained further results, and all of them form a relatively complete theory of formal feedback classification of nonlinear control systems. The first results of Kang and Krener were devoted to obtaining a normal form for single-input control-affine systems with controllable linear approximation and we will also start our systematic presentation in this section by discussing that case. A generalization to non-affine systems will be given at the end of this section while further developments (uncontrollable linear approximation and the problem of canonical forms) will be discussed in next sections.

This section is organized as follows. In Section 4.3.2, we will introduce the notation, used in the whole section as well as in Section 4.4 to Section 4.6. The main results are given in Section 4.3.3: a normal form for homogeneous systems, explicit transformations bringing to it, m -invariants, and normal

form under formal feedback. Finally, in Section 4.3.4, we will generalize the normal form to non-affine systems.

4.3.2 Notations

$P^{[m]}(\xi)$ denotes the space of homogeneous polynomials of degree m of the variables $\xi_1, \dots, \xi_n$; $P^{\leq m}(\xi)$ the space of polynomials of degree m of the variables $\xi_1, \dots, \xi_n$; and $P^{\geq m}(\xi)$ the space of formal power series of the variables $\xi_1, \dots, \xi_n$ starting from terms of degree m .

Analogously, $V^{[m]}(\xi)$ denotes the space of homogeneous vector fields whose components are in $P^{[m]}(\xi)$; $V^{\leq m}$ the space of polynomial vector fields whose components are in $P^{\leq m}(\xi)$; and $V^{\geq m}(\xi)$ the space of vector fields formal power series whose components are in $P^{\geq m}(\xi)$.

Notations $P^{[m]}(\xi, u)$, $V^{[m]}(\xi, u)$ represent, respectively, homogeneous polynomials and homogeneous polynomial vector fields depending on the state variables $\xi = (\xi_1, \dots, \xi_n)$ and control variable u , with homogeneity being understood with respect to the all variables (ξ, u) .

Because of various normal forms and various transformations that are used throughout the paper, we will maintain the following notation. Together with Σ , we will also consider its infinite Taylor series expansion Σ^∞ and its homogeneous part $\Sigma^{[m]}$ of degree m given, respectively, by the following systems

$$\begin{aligned}\Sigma^\infty : \dot{\xi} &= A\xi + Bu + \sum_{k=2}^{\infty} (f^{[k]}(\xi) + g^{[k-1]}(\xi)u) \\ \Sigma^{[m]} : \dot{\xi} &= A\xi + Bu + f^{[m]}(\xi) + g^{[m-1]}(\xi)u\end{aligned}$$

The systems Σ , $\Sigma^{[m]}$, and Σ^∞ will stand for the systems under consideration. Their state vector will be denoted by ξ and their control by u (x and v being used, respectively, for the state and control of various normal forms). The system $\Sigma^{[m]}$ (resp. Σ^∞) transformed via feedback will be denoted by $\tilde{\Sigma}^{[m]}$ (resp. $\tilde{\Sigma}^\infty$). Its state vector will be denoted by x , its control by v , and the vector fields, defining its dynamics, by $\tilde{f}^{[k]}$ and $\tilde{g}^{[k-1]}$. Feedback equivalence of homogeneous systems $\Sigma^{[m]}$ and $\tilde{\Sigma}^{[m]}$ will be established via a smooth feedback, specifically by homogeneous feedback Γ^m . On the other hand, feedback equivalence of systems Σ^∞ and $\tilde{\Sigma}^\infty$ will be established via a formal feedback Γ^∞ .

We will introduce two kinds of normal forms: Kang normal forms and dual normal forms (Section 4.3 and Section 4.5), as well as canonical forms and dual canonical forms (Section 4.4 and Section 4.6). The symbol “bar” will correspond to the vector field $\tilde{f}^{[m]}$ defining the Kang normal

forms $\Sigma_{\text{NF}}^{[m]}$ and $\Sigma_{\text{NF}}^{\infty}$ and the canonical form $\Sigma_{\text{CF}}^{\infty}$ as well as to the vector field $\bar{g}^{[m-1]}$ defining the dual normal forms $\Sigma_{\text{DNF}}^{[m]}$ and $\Sigma_{\text{DNF}}^{\infty}$ and the dual canonical form $\Sigma_{\text{DCF}}^{\infty}$. Analogously, the m -invariants (resp. dual m -invariants) of the system $\Sigma^{[m]}$ will be denoted by $a^{[m]j,i+2}$ (resp. $b_j^{[m-1]}$) and the m -invariants (resp. dual m -invariants) of the normal form $\Sigma_{\text{NF}}^{[m]}$ (resp. dual normal form $\Sigma_{\text{DNF}}^{[m]}$) by $\bar{a}^{[m]j,i+2}$ (resp. $\bar{b}_j^{[m-1]}$). Other normal forms will be discussed in Section 4.12.

4.3.3 Normal Form and m -Invariants

All objects, that is, functions, maps, vector fields, control systems, etc., are considered in a neighborhood of $0 \in \mathbb{R}^n$ and assumed to be C^{∞} -smooth. Let h be a smooth $\mathbb{R}$ -valued function. By

$$h(\xi) = h^{[0]}(\xi) + h^{[1]}(\xi) + h^{[2]}(\xi) + \cdots = \sum_{m=0}^{\infty} h^{[m]}(\xi)$$

we denote its infinite Taylor series expansion at $0 \in \mathbb{R}^n$, where $h^{[m]}(\xi)$ stands for a homogeneous polynomial of degree m .

Similarly, for a map ϕ of an open subset of $\mathbb{R}^n$ to $\mathbb{R}^n$ (resp. for a vector field f on an open subset of $\mathbb{R}^n$), we will denote by $\phi^{[m]}$ (resp. $f^{[m]}$) the homogeneous term of degree m of its Taylor series expansion at $0 \in \mathbb{R}^n$, that is, each component $\phi_j^{[m]}$ of $\phi^{[m]}$ (resp. $f_j^{[m]}$ of $f^{[m]}$) is a homogeneous polynomial of degree m in ξ .

Consider the Taylor series expansion of the system Σ given by

$$\Sigma^{\infty} : \dot{\xi} = F\xi + Gu + \sum_{m=2}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right) \quad (4.2)$$

where $F = (\partial f / \partial \xi)(0)$ and $G = g(0)$. Recall that we assume in this section that $f(0) = 0$ and $g(0) \neq 0$.

Consider also the Taylor series expansion Γ^{∞} of the feedback transformation Γ given by

$$\begin{aligned} x &= T\xi + \sum_{m=2}^{\infty} \phi^{[m]}(\xi) \\ \Gamma^{\infty} : \\ u &= K\xi + Lv + \sum_{m=2}^{\infty} \left(\alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v \right) \end{aligned} \quad (4.3)$$

where T is an invertible matrix and $L \neq 0$. Analogously to the Poincaré's approach presented in Section 4.2, we analyze the action of Γ^∞ on the system Σ^∞ step by step.

To start with, consider the linear system

$$\dot{\xi} = F\xi + Gu$$

Throughout the section we will assume that it is controllable. It can be thus transformed by a linear feedback transformation of the form

$$\Gamma^1 : \begin{aligned} x &= T\xi \\ u &= K\xi + Lv \end{aligned}$$

into the Brunovský canonical form (A, B) [49] and Example 1 in Section 4.1:

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Assuming that the linear part (F, G) , of the system Σ^∞ given by (4.2), has been transformed to the Brunovský canonical form (A, B) , we follow an idea of Kang and Krener [50, 54] and apply successively a series of transformations

$$\Gamma^m : \begin{aligned} x &= \xi + \phi^{[m]}(\xi) \\ u &= v + \alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v \end{aligned} \tag{4.4}$$

for $m = 2, 3, \dots$. A feedback transformation defined as an infinite series of successive compositions of Γ^m , $m = 1, 2, \dots$ is also denoted by Γ^∞ (i.e., $\Gamma^\infty = \dots \Gamma^m \circ \Gamma^{m-1} \circ \dots \circ \Gamma^1$) because, as a formal power series, it is of the form (4.3). We will not address the problem of convergence in general (see Section 4.9 and Section 4.13 for some comments on this issue and for a convergent class of analytic systems) and we will call such a series of successive compositions a *formal feedback transformation*.

Observe that each transformation Γ^m , for $m \geq 2$, leaves invariant all homogeneous terms of degree smaller than m of the system Σ^∞ and we will call Γ^m a *homogeneous feedback transformation of degree m* . We will study the action of Γ^m on the following *homogeneous system*

$$\Sigma^{[m]} : \dot{\xi} = A\xi + Bu + f^{[m]}(\xi) + g^{[m-1]}(\xi)u \tag{4.5}$$

Consider another homogeneous system $\tilde{\Sigma}^{[m]}$ given by

$$\tilde{\Sigma}^{[m]} : \dot{x} = Ax + Bv + \tilde{f}^{[m]}(x) + \tilde{g}^{[m-1]}(x)v \quad (4.6)$$

We will say that the homogeneous system $\Sigma^{[m]}$ is feedback equivalent to the homogenous system $\tilde{\Sigma}^{[m]}$ if there exists a homogeneous feedback transformation Γ^m of the form (4.4), which brings $\Sigma^{[m]}$ into $\tilde{\Sigma}^{[m]}$ modulo terms in $V^{\geq m+1}(x, v)$.

The starting point for formal classification of single-input control systems is the following result, proved by Kang [50].

PROPOSITION 2

The homogeneous feedback transformation Γ^m , defined by (4.4), brings the system $\Sigma^{[m]}$, given by (4.5), into $\tilde{\Sigma}^{[m]}$, given by (4.6), if and only if the following relations

$$\left\{ \begin{array}{l} L_{A\xi} \phi_j^{[m]} - \phi_{j+1}^{[m]}(\xi) = \tilde{f}_j^{[m]}(\xi) - f_j^{[m]}(\xi) \\ L_B \phi_j^{[m]}(\xi) = \tilde{g}_j^{[m-1]}(\xi) - g_j^{[m-1]}(\xi) \\ L_{A\xi} \phi_n^{[m]} + \alpha^{[m]}(\xi) = \tilde{f}_n^{[m]}(\xi) - f_n^{[m]}(\xi) \\ L_B \phi_n^{[m]}(\xi) + \beta^{[m-1]}(\xi) = \tilde{g}_n^{[m-1]}(\xi) - g_n^{[m-1]}(\xi) \end{array} \right. \quad (4.7)$$

hold for any $1 \leq j \leq n-1$, where $\phi_j^{[m]}$ are the components of $\phi^{[m]}$.

This proposition represents the essence of the method developed by Kang and Krener and has been used for many results in this chapter. The problem of studying the feedback equivalence of two control-affine systems Σ and $\tilde{\Sigma}$ requires, in general, solving the system (CDE) of first-order partial differential equations (as we have already explained in Section 4.1). On the other hand, if we perform the analysis step by step, then the problem of establishing the feedback equivalence of two systems $\Sigma^{[m]}$ and $\tilde{\Sigma}^{[m]}$ reduces to solving the algebraic system (4.7), called sometimes the control homological equation by its analogy with Poincaré's homological equation (HE) of Section 4.2. Indeed, (4.7) can be re-written in the following compact form

$$\begin{aligned} \text{ad}_{A\xi} \phi^{[m]}(\xi) &= \tilde{f}^{[m]}(\xi) - f^{[m]}(\xi) - B\alpha^{[m]}(\xi) \\ \text{ad}_B^{[m]}(\xi) &= \tilde{g}^{[m-1]}(\xi) - g^{[m-1]}(\xi) - B\beta^{[m-1]}(\xi) \end{aligned} \quad (\text{CHE})$$

which reduces to (HE) if the control vector field $B + g^{[m-1]}(\xi)$ is not present, with A playing the role of J . Therefore for control systems, solving the

differential equation (CDE) is replaced by an infinite sequence of algebraic homological equations (CHE) exactly like for dynamical systems, where the differential equation (DE) is replaced by an infinite sequence of homological equations (HE) (compare Section 4.2).

Using Proposition 2, Kang [50] proved the following:

THEOREM 3

The homogeneous system $\Sigma^{[m]}$ can be transformed, via a homogeneous feedback transformation Γ^m , into the following normal form

$$\Sigma_{\text{NF}}^{[m]} : \begin{cases} \dot{x}_1 = x_2 + \sum_{i=3}^n x_i^2 P_{1,i}^{[m-2]}(x_1, \dots, x_i) \\ \vdots \\ \dot{x}_j = x_{j+1} + \sum_{i=j+2}^n x_i^2 P_{j,i}^{[m-2]}(x_1, \dots, x_i) \\ \vdots \\ \dot{x}_{n-2} = x_{n-1} + x_n^2 P_{n-2,n}^{[m-2]}(x_1, \dots, x_n) \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = v \end{cases} \quad (4.8)$$

where $P_{j,i}^{[m-2]}(x_1, \dots, x_i)$ are homogeneous polynomials of degree $m - 2$ depending on the indicated variables.

To illustrate this result, consider the case $m = 2$, which actually was, for Kang and Krener [54], the starting point for the formal approach to feedback equivalence. Applying Theorem 3 to $m = 2$ yields that the homogeneous system $\Sigma^{[2]}$ can be transformed, via a homogeneous feedback transformation Γ^2 , into the following normal form:

$$\Sigma_{\text{NF}}^{[2]} : \begin{cases} \dot{x}_1 = x_2 + a_{1,3}x_3^2 + a_{1,4}x_4^2 + \dots + a_{1,n}x_n^2 \\ \dot{x}_2 = x_3 + a_{2,4}x_4^2 + \dots + a_{2,n}x_n^2 \\ \vdots \\ \dot{x}_{n-2} = x_{n-1} + a_{n-2,n}x_n^2 \\ \dot{x}_{n-1} = x_n + a_{n-1,n}x_n^2 \\ \dot{x}_n = u \end{cases}$$

where $a_{j,i} \in \mathbb{R}$. Notice that the general normal form $\Sigma_{\text{NF}}^{[m]}$ exhibits the same triangular triangular structure as $\Sigma_{\text{NF}}^{[2]}$, the only difference being the replacement of the constants $a_{j,i}$ by the polynomials $P_{j,i}^{[m-2]}(x_1, \dots, x_i)$.

Now we will compute the number of constants $a_{j,i}$ (for $m = 2$) and that of the polynomials (in the general case) present in the normal forms. Compare the analysis given subsequently (performed for the homogeneous system $\Sigma^{[m]}$) with a similar analysis given for general multi-input systems Π and control-affine systems Σ in Section 4.1. Recall that the dimension of the space of polynomials $P^{[m]}$ of degree m of n variables and of the space $V^{[m]}$ of polynomial vector fields on $\mathbb{R}^n$, all of whose components belong to $P^{[m]}$, are, respectively

$$\frac{(n+m-1)!}{m!(n-1)!} \quad \text{and} \quad n \frac{(n+m-1)!}{m!(n-1)!}$$

Homogeneous systems $\Sigma^{[m]}$ are given by two vector fields $f^{[m]} \in V^{[m]}$ and $g^{[m-1]} \in V^{[m-1]}$. Therefore, the dimension of the space of single-input systems, homogenous of degree m , is

$$d_{\Sigma^{[m]}} = n \frac{(n+m-1)!}{m!(n-1)!} + n \frac{(n+m-2)!}{(m-1)!(n-1)!}$$

The feedback group Γ^m is given by n components of the diffeomorphism $\phi^{[m]}$, each in $P^{[m]}$, and two functions $\alpha^{[m]} \in P^{[m]}$ and $\beta^{[m-1]} \in P^{[m-1]}$. Hence the dimension of Γ^m is

$$d_{\Gamma^m} = n \frac{(n+m-1)!}{m!(n-1)!} + \frac{(n+m-1)!}{m!(n-1)!} + \frac{(n+m-2)!}{(m-1)!(n-1)!}$$

Both dimensions are polynomials of degree $n-1$ of m and their difference is thus also a polynomial of degree $n-1$ of m starting with

$$d_{\Sigma^{[m]}} - d_{\Gamma^m} = \frac{n-2}{(n-1)!} m^{n-1} + \dots,$$

where dots stand for lower order terms. Observe that the dimension of the space of $n-2$ functions, each belonging to $P^{[m]}$, is also a polynomial of degree $n-1$ of m starting with

$$(n-2) \frac{(n+m-1)!}{m!(n-1)!} = \frac{n-2}{(n-1)!} m^{n-1} + \dots$$

which explains why in the normal form $\Sigma_{\text{NF}}^{[m]}$ we have $n-2$ polynomials of n variables. Since

$$d_{\Sigma^{[m]}} - d_{\Gamma^m} < (n-2) \frac{(n+m-1)!}{m!(n-1)!}$$

it follows that polynomials of fewer variables show up in the normal form $\Sigma_{\text{NF}}^{[m]}$. An analogous argument applied to m tending to infinity explains the appearance of $n - 2$ functions of n variables in the normal form $\Sigma_{\text{NF}}^{\infty}$ (see Theorem 6). To calculate the exact number of invariants in the form $\Sigma^{[m]}$ (which is bounded from below by $d_{\Sigma^{[m]}} - d_{\Gamma^m}$), we have to study the action of Γ^m on the space of homogeneous systems of degree m . This action is not free, the isotropy group being of dimension 1 [50, 83] (see also Proposition 3 for a detailed calculation). This can be illustrated by the homogeneous system $\Sigma^{[2]}$ of degree 2, for which $d_{\Sigma^{[2]}} = n(n(n+1))/2 + mn$ (we have n components of $f^{[2]}$ and n components of $g^{[1]}$) and $d_{\Gamma^2} = n(n(n+1))/2 + (n(n+1))/2 + n$ (we have n components of $\phi^{[2]}$ and the function $\alpha^{[2]}$ and $\beta^{[1]}$). It follows that $d_{\Sigma^{[2]}} - d_{\Gamma^2} = (n^2 - 3n)/2$ while the number of parameters $a_{j,i}$ (which is actually the number of invariants of $\Sigma^{[2]}$, see the next section) is $((n-1)(n-2))/2$. The difference $((n-1)(n-2))/2 - (n^2 - 3n)/2 = 1$ is actually the dimension of the isotropy subgroup of Γ^2 , which is the dimension of the group of symmetries of any $\Sigma^{[2]}$ (see Section 4.11).

The two following questions concerning the normal form $\Sigma_{\text{NF}}^{[m]}$ are important and arise naturally:

1. Are the polynomials $P_{j,i}^{[m-2]}$ invariant, that is, unique under feedback Γ^m ?
2. How to bring a given system $\Sigma^{[m]}$ into its normal form $\Sigma_{\text{NF}}^{[m]}$?

The answer to question 1 is positive, and to construct invariants under homogeneous feedback transformations, define the vector fields

$$X_i^{m-1}(\xi) = (-1)^i \text{ad}_{A\xi + f^{[m]}(\xi)}^i (B + g^{[m-1]}(\xi))$$

and let $X_i^{[m-1]}$ be its homogeneous part of degree $m-1$. By π_i we will denote the projection on the subspace

$$W_i = \{\xi = (\xi_1, \dots, \xi_n)^T \in R^n : \xi_{i+1} = \dots = \xi_n = 0\}$$

that is

$$\pi_i(\xi) = (\xi_1, \dots, \xi_i, 0, \dots, 0)$$

Following Kang [50], we denote by $a^{[m]j,i+2}(\xi)$ the homogeneous part of degree $m-2$ of the polynomials

$$\begin{aligned} CA^{j-1} \left[X_i^{m-1}, X_{i+1}^{m-1} \right] (\pi_{n-i}(\xi)) &= CA^{j-1} \\ &\quad \times \left(\text{ad}_{A^i B} X_{i+1}^{[m-1]} - \text{ad}_{A^{i+1} B} X_i^{[m-1]} \right) (\pi_{n-i}(\xi)) \end{aligned}$$

where $C = (1, 0, \dots, 0)^T \in \mathbb{R}^n$ and $(j, i) \in \Delta \subset \mathbb{N} \times \mathbb{N}$, defined by

$$\Delta = \{(j, i) \in \mathbb{N} \times \mathbb{N} : 1 \leq j \leq n-2, 0 \leq i \leq n-j-2\}$$

The homogeneous polynomials $a^{[m]j,i+2}$, for $(j, i) \in \Delta$, will be called *m-invariants* of $\Sigma^{[m]}$, under the action of Γ^m .

The following result of Kang [50] asserts that *m-invariants* $a^{[m]j,i+2}$, for $(j, i) \in \Delta$, are complete invariants of homogeneous feedback and, moreover, illustrates their meaning for the homogeneous normal form $\Sigma_{\text{NF}}^{[m]}$.

Consider two homogeneous systems $\Sigma^{[m]}$ and $\tilde{\Sigma}^{[m]}$ and let

$$\{a^{[m]j,i+2} : (j, i) \in \Delta\}, \quad \text{and} \quad \{\tilde{a}^{[m]j,i+2} : (j, i) \in \Delta\}$$

denote, respectively, their *m-invariants*. The following result was proved by Kang [50]:

THEOREM 4

The m-invariants have the following properties:

1. Two homogeneous systems $\Sigma^{[m]}$ and $\tilde{\Sigma}^{[m]}$ are equivalent via a homogeneous feedback transformation Γ^m if and only if

$$a^{[m]j,i+2} = \tilde{a}^{[m]j,i+2}, \quad \text{for any } (j, i) \in \Delta$$

2. The *m-invariants* $\tilde{a}^{[m]j,i+2}$ of the normal form $\Sigma_{\text{NF}}^{[m]}$, defined by (4.8), are given by

$$\tilde{a}^{[m]j,i+2}(x) = \frac{\partial^2}{\partial x_{n-i}^2} x_{n-i}^2 P_{j,n-i}^{[m-2]}(x_1, \dots, x_{n-i}), \quad \text{for any } (j, i) \in \Delta \quad (4.9)$$

To answer question 2, we will construct an explicit feedback transformation that brings the homogeneous system $\Sigma^{[m]}$ to its normal form $\Sigma_{\text{NF}}^{[m]}$. Define the homogeneous polynomials $\psi_{j,i}^{[m-1]}(\xi)$ by setting $\psi_{j,0}^{[m-1]}(\xi) = \psi_{1,1}^{[m-1]}(\xi) = 0$,

$$\psi_{j,i}^{[m-1]}(\xi) = -CA^{j-1} \left(\text{ad}_{A\xi}^{n-i} g^{[m-1]} + \sum_{t=1}^{n-i} (-1)^t \text{ad}_{A\xi}^{t-1} \text{ad}_{A^{n-i-t}B} f^{[m]} \right)$$

if $1 \leq j < i \leq n$ and

$$\begin{aligned} \psi_{j,i}^{[m-1]}(\xi) &= L_{A^{n-i}B} f_{j-1}^{[m]}(\pi_i(\xi)) + L_{A\xi} \psi_{j-1,i}^{[m-1]}(\pi_i(\xi)) \\ &\quad + \psi_{j-1,i-1}^{[m-1]}(\pi_{i-1}(\xi)) + \int_0^{\xi_i} L_{A^{n-i+1}B} \psi_{j-1,i}^{[m-1]}(\pi_i(\xi)) d\xi_i \end{aligned} \quad (4.10)$$

if $1 \leq i \leq j$, where $\psi_{j,i}^{[m-1]}(\pi_i(\xi))$ is the restriction of $\psi_{j,i}^{[m-1]}(\xi)$ to the subspace W_i . Define the components $\phi_j^{[m]}$ of $\phi^{[m]}$, for $1 \leq j \leq n$, and the feedback $(\alpha^{[m]}, \beta^{[m-1]})$ by

$$\begin{aligned} \phi_j^{[m]}(\xi) &= \sum_{i=1}^n \int_0^{\xi_i} \psi_{j,i}^{[m-1]}(\pi_i(\xi)) d\xi_i, \quad 1 \leq j \leq n-1 \\ \phi_n^{[m]}(\xi) &= f_{n-1}^{[m]}(\xi) + L_{A\xi} \phi_{n-1}^{[m]}(\xi) \\ \alpha^{[m]}(\xi) &= - \left(f_n^{[m]}(\xi) + L_{A\xi} \phi_n^{[m]}(\xi) \right) \\ \beta^{[m-1]}(\xi) &= - \left(g_n^{[m-1]}(\xi) + L_B \phi_n^{[m]}(\xi) \right) \end{aligned} \quad (4.11)$$

We have the following result [83]:

THEOREM 5

The homogeneous feedback transformation

$$\begin{aligned} \Gamma^m : \quad & x = \xi + \phi^{[m]}(\xi) \\ & u = v + \alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v \end{aligned}$$

where $\alpha^{[m]}$, $\beta^{[m-1]}$, and the components $\phi_j^{[m]}$ of $\phi^{[m]}$ are defined by (4.11), brings the homogeneous system $\Sigma^{[m]}$ into its normal form $\Sigma_{NF}^{[m]}$ given by (4.8).

Example 2

To illustrate the results of this section, we consider the system $\Sigma^{[m]}$, given by (4.5) on $\mathbb{R}^3$. Theorem 3 implies that the system $\Sigma^{[m]}$ is equivalent, via a homogeneous feedback transformation Γ^m defined by (4.11), to its normal form $\Sigma_{NF}^{[m]}$ (see (4.8))

$$\begin{aligned} \dot{x}_1 &= x_2 + x_3^2 p^{[m-2]}(x_1, x_2, x_3) \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= v \end{aligned}$$

where $p^{[m-2]}(x_1, x_2, x_3)$ is a homogeneous polynomial of degree $m-2$ of the variables x_1, x_2, x_3 .

As we have already mentioned, Poincaré's method allows to replace the partial differential equation (CDE) (given in Section 4.1) by solving successively linear algebraic equations defined by the control homological equation (CHE) [50, 54] and Proposition 2. The solvability of this equation was proved earlier [50, 54] while Theorem 5 provides an explicit solution (in the form of the transformations (4.11) that are easily computable via differentiation and integration of homogeneous polynomials) to the control homological equation. Consequently, for any given control system, Theorem 5 gives transformations bringing the homogeneous part of the system into its normal form. For example, if the system is feedback linearizable, up to order $m_0 - 1$ [56], then a diffeomorphism and a feedback compensating all nonlinearities of degree lower than m_0 can be calculated explicitly without solving partial differential equations (compare Section 4.9). More generally, by a successive application of transformations given by (4.11) we can bring the system, without solving partial differential equations, to its normal form given in Theorem 6.

Consider the system Σ^∞ of the form (4.2) and recall that we assume the linear part (F, G) to be controllable. Apply successively to Σ^∞ a series of transformations $\Gamma^m, m = 1, 2, 3, \dots$, such that each Γ^m brings $\Sigma^{[m]}$ to its normal form $\Sigma_{\text{NF}}^{[m]}$. More precisely, bring (F, G) into the Brunonvský canonical form (A, B) via a linear feedback Γ^1 and denote $\Sigma^{\infty,1} = \Gamma_*^1(\Sigma^\infty)$. Assume that a system $\Sigma^{\infty,m-1}$ has been defined. Let Γ^m be a homogeneous feedback transformation transforming $\Sigma^{[m]}$, which is the homogeneous part of degree m of $\Sigma^{\infty,m-1}$, to the normal form $\Sigma_{\text{NF}}^{[m]}$ (Γ^m can be taken, for instance, as the transformations defined by (4.11)). Define $\Sigma^{\infty,m} = \Gamma_*^m(\Sigma^{\infty,m-1})$. Notice that we apply Γ^m to the whole system $\Sigma^{\infty,m-1}$ (and not only to its homogeneous part $\Sigma^{[m]}$). Successive iteration of Theorem 3 gives the following result of Kang [50].

THEOREM 6

There exists a formal feedback transformation Γ^∞ which brings the system Σ^∞ to a normal form $\Sigma_{\text{NF}}^\infty$ given by

$$\Sigma_{\text{NF}}^\infty : \begin{cases} \dot{x}_1 = x_2 + \sum_{i=3}^n x_i^2 P_{1,i}(x_1, \dots, x_i) \\ \vdots \\ \dot{x}_j = x_{j+1} + \sum_{i=j+2}^n x_i^2 P_{j,i}(x_1, \dots, x_i) \\ \vdots \\ \dot{x}_{n-2} = x_{n-1} + x_n^2 P_{n-2,n}(x_1, \dots, x_n) \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = v \end{cases} \quad (4.12)$$

where $P_{j,i}(x_1, \dots, x_i)$ are formal power series depending on the indicated variables.

Example 3

Consider a system Σ defined on $\mathbb{R}^3$ whose linear part is controllable (compare Example 2). Theorem 6 implies that the system Σ is equivalent, via a formal feedback transformation Γ^∞ , to its normal form $\Sigma_{\text{NF}}^\infty$

$$\begin{aligned}\dot{x}_1 &= x_2 + x_3^2 P(x_1, x_2, x_3) \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= v\end{aligned}$$

where $P(x_1, x_2, x_3)$ is a formal power series of the variables x_1, x_2, x_3 .

4.3.4 Normal Form for Non-affine Systems

In this section, we will generalize normal forms $\Sigma_{\text{NF}}^{[m]}$ and $\Sigma_{\text{NF}}^\infty$ to general systems. As we explained in Section 4.1, such a generalization can be performed using Proposition 1.

Consider a general control system of the form

$$\Pi : \dot{\xi} = F(\xi, u)$$

around an equilibrium point (ξ_0, u_0) , that is, $F(\xi_0, u_0) = 0$. Without loss of generality, we can assume that $(\xi_0, u_0) = (0, 0)$. Together with Π we will consider its infinite Taylor series expansion

$$\Pi^\infty : \dot{\xi} = F\xi + Gu + \sum_{m=2}^{\infty} F^{[m]}(\xi, u)$$

where $F^{[m]}(\xi, u)$ stands for homogeneous terms of degree m and homogeneity is understood in this section with respect to the state and control variables together.

Consider the feedback transformation Υ (compare Section 4.1)

$$\begin{aligned}\Upsilon : \quad x &= \phi(\xi) \\ v &= \psi(\xi, u)\end{aligned}$$

and its Taylor series expansion Υ^∞ given by

$$\begin{aligned} x &= T\xi + \sum_{m=2}^{\infty} \phi^{[m]}(\xi) \\ \Upsilon^\infty : \\ v &= K\xi + Lu + \sum_{m=2}^{\infty} \psi^{[m]}(\xi, u) \end{aligned}$$

where T is an invertible matrix and $L \neq 0$.

We will assume throughout this section that the pair (F, G) is controllable and so we can suppose that it is in the Brunovský canonical form (A, B) . Like in the control-affine case, we will consider the action of the homogenous part Υ^m of Υ^∞ given by

$$\begin{aligned} \Upsilon^m : \quad x &= \xi + \phi^{[m]}(\xi) \\ v &= u + \psi^{[m]}(\xi, u) \end{aligned}$$

on the homogeneous part $\Pi^{[m]}$ of Π^∞ given by

$$\Pi^{[m]} : \dot{\xi} = A\xi + Bu + F^{[m]}(\xi, u)$$

Combining Theorem 3 with Proposition 1 leads to the following result:

THEOREM 7

The general homogeneous system $\Pi^{[m]}$ can be transformed, via a homogeneous feedback transformation Υ^m , into the following normal form

$$\Pi_{\text{NF}}^{[m]} : \begin{cases} \dot{x}_1 = x_2 + \sum_{i=3}^n x_i^2 P_{1,i}^{[m-2]}(x_1, \dots, x_i) + v^2 P_1^{[m-2]}(x_1, \dots, x_n, v) \\ \vdots \\ \dot{x}_j = x_{j+1} + \sum_{i=j+2}^n x_i^2 P_{j,i}^{[m-2]}(x_1, \dots, x_i) + v^2 P_j^{[m-2]}(x_1, \dots, x_n, v) \\ \vdots \\ \dot{x}_{n-2} = x_{n-1} + x_n^2 P_{n-2,n}^{[m-2]}(x_1, \dots, x_n) + v^2 P_{n-1}^{[m-2]}(x_1, \dots, x_n, v) \\ \dot{x}_{n-1} = x_n + v^2 P_n^{[m-2]}(x_1, \dots, x_n, v) \\ \dot{x}_n = v \end{cases} \quad (4.13)$$

where $P_{j,i}^{[m-2]}(x_1, \dots, x_i)$ and $P_j^{[m-2]}(x_1, \dots, x_n, v)$ are homogeneous polynomials of degree $m-2$ depending on the indicated variables.

Notice that formally the aforementioned normal form can be obtained as follows. Consider the affine normal form $\Sigma_{\text{NF}}^{[m]}$ and apply the reduction defined as the inverse of the extension (preintegration) described just before Proposition 1. More precisely, assume that $\Sigma_{\text{NF}}^{[m]}$ is controlled by x_n , so skip the last equation $\dot{x}_n = v$ and denote x_n by v . What we obtain is an $(n - 1)$ -dimensional system, nonlinear with respect to v , which actually gives the $(n - 1)$ -dimensional form $\Pi_{\text{NF}}^{[m]}$.

Like in the control-affine case, a successive application of Theorem 7 gives a formal normal form for Π_{NF}^{∞} under Υ^{∞} . It has the same structure as $\Pi_{\text{NF}}^{[m]}$, the only difference being that the polynomials $P_{j,i}^{[m-2]}(x_1, \dots, x_i)$ and $P_j^{[m-2]}(x_1, \dots, x_n, v)$ are replaced by formal power series of the same variables.

We will end up with a simple example, which, actually, is a nonaffine version of Example 2 and Example 3.

Example 4

Consider the general system $\Pi^{[m]}$ on $\mathbb{R}^2$. Theorem 6 implies that the system $\Pi^{[m]}$ is equivalent, via a homogeneous feedback transformation Υ^m to its normal form $\Pi_{\text{NF}}^{[m]}$, see (4.13):

$$\begin{aligned}\dot{x}_1 &= x_2 + v^2 P_1^{[m-2]}(x_1, x_2, v) \\ \dot{x}_2 &= v\end{aligned}$$

where $P_1^{[m-2]}(x_1, x_2, v)$ is a homogeneous polynomial of degree $m - 2$ of the variables x_1, x_2 , and v .

Consequently, the general system Π^{∞} on $\mathbb{R}^2$ is equivalent, via a formal feedback transformation Υ^{∞} to its normal form Π_{NF}^{∞} :

$$\begin{aligned}\dot{x}_1 &= x_2 + v^2 P_1(x_1, x_2, v) \\ \dot{x}_2 &= v\end{aligned}$$

where $P_1(x_1, x_2, v)$ is a formal power series of the variables x_1, x_2 , and v .

4.4 Canonical Form for Single-Input Systems with Controllable Linearization

As proved by Kang and recalled in Theorem 4, the normal form $\Sigma_{\text{NF}}^{[m]}$ is unique under homogeneous feedback transformation Γ^m . The normal form

$\Sigma_{\text{NF}}^\infty$ is constructed by a successive application of homogeneous transformations Γ^m , for $m \geq 1$, which bring the corresponding homogeneous systems $\Sigma^{[m]}$ into their normal forms $\Sigma_{\text{NF}}^{[m]}$. Therefore, a natural and fundamental question which arises is whether the system Σ^∞ can admit two different normal forms, that is, whether the normal forms given by Theorem 6 are in fact canonical forms under a general formal feedback transformations of the form Γ^∞ . It turns out that a given system can admit different normal forms, as shown in the following example of Kang [50]. The main reason for the nonuniqueness of the normal form $\Sigma_{\text{NF}}^\infty$ is that, although the normal form $\Sigma_{\text{NF}}^{[m]}$ is unique, homogeneous feedback transformation Γ^m bringing $\Sigma^{[m]}$ into $\Sigma_{\text{NF}}^{[m]}$ is not. It is this small group of homogeneous feedback transformations of order m that preserve $\Sigma_{\text{NF}}^{[m]}$ (described by Proposition 3), which causes the nonuniqueness of $\Sigma_{\text{NF}}^\infty$.

Example 5

Consider the following system

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 + \xi_3^2 - 2\xi_1\xi_3^2 \\ \dot{\xi}_2 &= \xi_3 \\ \dot{\xi}_3 &= u\end{aligned}\tag{4.14}$$

on $\mathbb{R}^3$. Clearly, this system is in Kang normal form (compare with Theorem 6), say $\Sigma_{1,\text{NF}}^\infty$. The feedback transformation

$$\begin{aligned}\Gamma^{\leq 3}: \quad x_1 &= \xi_1 - \xi_1^2 - \frac{4}{3}\xi_2^3 \\ x_2 &= \xi_2 - 2\xi_1\xi_2 \\ x_3 &= \xi_3 - 2\left(\xi_2^2 + \xi_1\xi_3\right) - 2\xi_2\xi_3^2 \\ u &= v + 6\xi_2\xi_3 + 12\xi_1\xi_2\xi_3 - 4\xi_3^3 + 2\left(\xi_1 + 2\xi_1^2 + 2\xi_2\xi_3\right)v\end{aligned}$$

brings the system (4.14) into the form

$$\begin{aligned}\dot{x}_1 &= x_2 + x_3^2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= v\end{aligned}$$

modulo terms in $V^{\geq 4}(x, v)$. Applying successively homogeneous feedback transformations Γ^m given, for any $m \geq 4$, by (4.11), we transform the

aforementioned system into the following normal form $\Sigma_{2,\text{NF}}^\infty$:

$$\begin{aligned}\dot{x}_1 &= x_2 + x_3^2 + x_3^2 P(x) \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= v\end{aligned}\tag{4.15}$$

where P is a formal power series whose 1-jet at $0 \in \mathbb{R}^3$ vanishes. The systems (4.14) and (4.15) are in their normal forms ($\Sigma_{1,\text{NF}}^\infty$ and $\Sigma_{2,\text{NF}}^\infty$, respectively) and, moreover, the systems are feedback equivalent, but the system (4.15) does not contain any term of degree 3. As a consequence, the normal form $\Sigma_{\text{NF}}^\infty$ is not unique under feedback transformations.

A natural and important problem is thus to construct a canonical form and the aim of this section is indeed to construct a canonical form for Σ^∞ under feedback transformation Γ^∞ .

Consider the system Σ^∞ of the form

$$\Sigma^\infty: \dot{\xi} = F\xi + Gu + \sum_{m=2}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right)\tag{4.16}$$

Since its linear part (F, G) is assumed to be controllable, we bring it, via a linear transformation and linear feedback, to the Brunovský canonical form (A, B) . Let the first homogeneous term of Σ^∞ which cannot be annihilated by a feedback transformation be of degree m_0 . As proved by Krener [56], the degree m_0 is given by the largest integer such that all distributions $\mathcal{D}^k = \text{span}\{g, \dots, \text{ad}_f^{k-1}g\}$, for $1 \leq k \leq n-1$, are involutive modulo terms of order $m_0 - 2$. We can thus, due to Theorem 3 and Theorem 4, assume that, after applying a suitable feedback $\Gamma^{\leq m_0}$, the system Σ^∞ takes the form

$$\dot{\xi} = A\xi + Bu + \bar{f}^{[m_0]}(\xi) + \sum_{m=m_0+1}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right)$$

where (A, B) is in Brunovský canonical form and the first nonvanishing homogeneous vector field $\bar{f}^{[m_0]}$ is in the normal form (by Theorem 3) with components given by

$$\bar{f}_j^{[m_0]}(\xi) = \begin{cases} \sum_{i=j+2}^n \xi_i^2 P_{j,i}^{[m_0-2]}(\xi_1, \dots, \xi_i), & 1 \leq j \leq n-2 \\ 0, & n-1 \leq j \leq n \end{cases}$$

Let $(i_1, \dots, i_{n-s})$, where $i_1 + \dots + i_{n-s} = m_0$ and $i_{n-s} \geq 2$, be the largest, in the lexicographic ordering, $(n-s)$ -tuple of nonnegative integers such that

for some $1 \leq j \leq n-2$, we have

$$\frac{\partial^{m_0} \bar{f}_j^{[m_0]}}{\partial \xi_1^{i_1} \cdots \partial \xi_{n-s}^{i_{n-s}}} \neq 0$$

Define

$$j^* = \sup \left\{ j = 1, \dots, n-2 : \frac{\partial^{m_0} \bar{f}_j^{[m_0]}}{\partial \xi_1^{i_1} \cdots \partial \xi_{n-s}^{i_{n-s}}} \neq 0 \right\}$$

The following results, whose proofs are detailed elsewhere [83], describe the canonical form obtained by the authors.

THEOREM 8

The system Σ^∞ given by (4.16) is equivalent by a formal feedback Γ^∞ to a system of the form

$$\Sigma_{\text{CF}}^\infty : \dot{x} = Ax + Bv + \sum_{m=m_0}^{\infty} \bar{f}^{[m]}(x)$$

where, for any $m \geq m_0$, the components $\bar{f}_j^{[m]}(x)$ of $\bar{f}^{[m]}(x)$ are given by

$$\bar{f}_j^{[m]}(x) = \begin{cases} \sum_{i=j+2}^n x_i^2 P_{j,i}^{[m-2]}(x_1, \dots, x_i), & 1 \leq j \leq n-2 \\ 0, & n-1 \leq j \leq n \end{cases} \quad (4.17)$$

additionally, we have

$$\frac{\partial^{m_0} \bar{f}_{j^*}^{[m_0]}}{\partial x_1^{i_1} \cdots \partial x_{n-s}^{i_{n-s}}} = \pm 1 \quad (4.18)$$

and, moreover, for any $m \geq m_0 + 1$

$$\frac{\partial^{m_0} \bar{f}_{j^*}^{[m]}}{\partial x_1^{i_1} \cdots \partial x_{n-s}^{i_{n-s}}}(x_1, 0, \dots, 0) = 0. \quad (4.19)$$

The form $\Sigma_{\text{CF}}^\infty$ satisfying (4.17)–(4.19) will be called the canonical form of Σ^∞ . The name is justified by the following theorem.

THEOREM 9

Two systems Σ_1^∞ and Σ_2^∞ are formally feedback equivalent if and only if their canonical forms $\Sigma_{1,\text{CF}}^\infty$ and $\Sigma_{2,\text{CF}}^\infty$ coincide.

Kang [50], generalizing [54], proved that any system Σ^∞ can be brought by a formal feedback into the normal form $\Sigma_{\text{NF}}^\infty$, for which (4.17) is satisfied. He also observed that his normal forms are not unique (see Example 5). Our results (Theorem 8 and Theorem 9) complete his study. We show that for each degree m of homogeneity we can use a one-dimensional subgroup of feedback transformations which preserves the “triangular” structure of (4.17) and at the same time allows us to normalize one higher order term. The form of (4.18) and (4.19) is a result of this normalization. These one-dimensional subgroups of feedback transformations are given by the following proposition.

PROPOSITION 3

The transformation Γ^m given by (4.4) leaves the system $\Sigma^{[m]}$ defined by (4.5) invariant if and only if

$$\begin{aligned}\phi_j^{[m]} &= a_m L_{A\xi}^{j-1} \xi_1^m, \quad 1 \leq j \leq n \\ \alpha^{[m]} &= -a_m L_{A\xi}^n \xi_1^m \\ \beta^{[m-1]} &= -a_m L_B L_{A\xi}^{n-1} \xi_1^m\end{aligned}\tag{4.20}$$

where a_m is an arbitrary real parameter.

Theorem 8 establishes an equivalence of the system Σ^∞ with its canonical form $\Sigma_{\text{CF}}^\infty$ via a formal feedback. Its direct corollary yields the following result for equivalence under a smooth feedback of the form

$$\begin{aligned}x &= \phi(\xi) \\ \Gamma : \\ u &= \alpha(\xi) + \beta(\xi)v\end{aligned}$$

up to an arbitrary order. Indeed, we have the following:

COROLLARY 2

Consider a smooth control system

$$\Sigma : \dot{\xi} = f(\xi) + g(\xi)u$$

For any positive integer k we have:

1. *There exists a smooth feedback Γ transforming Σ , locally around $0 \in \mathbb{R}^n$, into its canonical form $\Sigma_{\text{CF}}^{\leq k}$ given by:*

$$\Sigma_{\text{CF}}^{\leq k} : \dot{x} = Ax + Bv + \sum_{m=m_0}^k \bar{f}^{[m]}(x)$$

modulo terms in $V^{\geq k+1}(x, v)$, where the components $\tilde{f}_j^{[m]}(x)$ of $\tilde{f}^{[m]}(x)$, for any $m_0 \leq m \leq k$, satisfy (4.17)–(4.19).

2. Feedback equivalence of Σ and $\Sigma_{\text{CF}}^{\leq k}$, modulo terms in $V^{\geq k+1}(x, v)$, can be established via a polynomial feedback transformation $\Gamma^{\leq k}$ of degree k .
3. Two smooth systems Σ_1 and Σ_2 are feedback equivalent modulo terms in $V^{\geq k+1}(x, v)$ if and only if their canonical forms $\Sigma_{1,\text{CF}}^{\leq k}$ and $\Sigma_{2,\text{CF}}^{\leq k}$ coincide.

This corollary follows directly from Theorem 8 and Theorem 9.

To end this section will illustrate our results by two examples.

Example 6

Let us reconsider the system Σ given by Example 3. It is equivalent, via a formal feedback, to the normal form

$$\begin{aligned}\dot{x}_1 &= x_2 + x_3^2 P(x_1, x_2, x_3) \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= v\end{aligned}$$

where $P(x_1, x_2, x_3)$ is a formal power series. Assume, for simplicity, that $m_0 = 2$, which is equivalent to the following generic condition: g , $\text{ad}_f g$, and $[g, \text{ad}_f g]$ are linearly independent at $0 \in \mathbb{R}^3$. This implies that we can express $P = P(x_1, x_2, x_3)$ as

$$P = c + P_1(x_1) + x_2 P_2(x_1, x_2) + x_3 P_3(x_1, x_2, x_3)$$

where $c \neq 0$ and $P_1(0) = 0$. Observe that any $P(x_1, x_2, x_3)$ of the earlier form gives a normal form $\Sigma_{\text{NF}}^\infty$. To get the canonical form $\Sigma_{\text{CF}}^\infty$, we use Theorem 8 which assures the existence of a feedback transformation Γ^∞ of the form

$$\begin{aligned}\tilde{x} &= \phi(x) \\ v &= \alpha(x) + \beta(x)\tilde{v}\end{aligned}$$

which normalizes the constant c and annihilates the formal power series $P_1(x_1)$. More precisely, Γ^∞ transforms Σ into its canonical form $\Sigma_{\text{CF}}^\infty$

$$\begin{aligned}\dot{\tilde{x}}_1 &= \tilde{x}_2 + \tilde{x}_3^2 \tilde{P}(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3) \\ \dot{\tilde{x}}_2 &= \tilde{x}_3 \\ \dot{\tilde{x}}_3 &= \tilde{v}\end{aligned}$$

where the formal power series $\tilde{P}(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$ is of the form

$$\tilde{P}(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3) = 1 + \tilde{x}_2 \tilde{P}_2(\tilde{x}_1, \tilde{x}_2) + \tilde{x}_3 \tilde{P}_3(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$$

clearly showing a difference between the normal and canonical form: in the latter, the free term is normalized and the term depending on x_1 is annihilated.

Now, we give an example of constructing the canonical form for a physical model of a variable length pendulum.

Example 7

Consider the variable length pendulum of Bressan and Rampazzo [10] (see also [19]). We denote by ξ_1 the length of the pendulum, by ξ_2 its velocity, by ξ_3 the angle with respect to the horizontal, and by ξ_4 the angular velocity. The control $u = \dot{\xi}_4 = \ddot{\xi}_3$ is the angular acceleration. The mass is normalized to 1. The equations are [10, 19]:

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= -g \sin \xi_3 + \xi_1 \xi_4^2 \\ \dot{\xi}_3 &= \xi_4 \\ \dot{\xi}_4 &= u\end{aligned}$$

where g denotes the gravity. Note that if we suppose to control the angular velocity $\xi_4 = \dot{\xi}_3$, which is the case of Refs. [10, 19], then the system is three-dimensional but the control enters nonlinearly.

At any equilibrium point $\xi_0 = (\xi_{10}, \xi_{20}, \xi_{30}, \xi_{40})^T = (\xi_{10}, 0, 0, 0)^T$, the linear part of the system is controllable. Our goal is to produce, for the variable length pendulum, a normal form and the canonical form as well as to answer the question regarding whether the systems corresponding to various values of the gravity constant g are feedback equivalent. To get a normal form, put

$$\begin{aligned}x_1 &= \xi_1 \\ x_2 &= \xi_2 \\ x_3 &= -g \sin \xi_3 \\ x_4 &= -g \xi_4 \cos \xi_3 \\ v &= g \xi_4^2 \sin \xi_3 - u g \cos \xi_3\end{aligned}$$

The system becomes

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 + x_4^2 \frac{x_1}{g^2 - x_3^2} \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= v\end{aligned}$$

which gives a normal form. Indeed, we rediscover $\Sigma_{\text{NF}}^\infty$, given by (4.12), with $P_{1,3} = 0$, $P_{1,4} = 0$, and

$$P_{2,4} = \frac{x_1}{g^2 - x_3^2}$$

To bring the system to its canonical form $\Sigma_{\text{CF}}^\infty$, first observe that $m_0 = 3$. Indeed, the function $x_4^2(x_1/(g^2 - x_3^2))$ starts with third-order terms, which corresponds to the fact that the invariants $a^{[2]j,i+2}$ vanish for any $1 \leq j \leq 2$ and any $0 \leq i \leq 2 - j$. The only nonzero component of $\bar{f}^{[3]}$ is $\bar{f}_2^{[3]} = x_4^2 P_{2,4}^{[1]}$. Hence $j^* = 2$ and the only, and thus the largest, quadruplet (i_1, i_2, i_3, i_4) of nonnegative integers, satisfying $i_1 + i_2 + i_3 + i_4 = 3$ and such that

$$\frac{\partial^3 \bar{f}_2^{[3]}}{\partial x_1^{i_1} \cdots \partial x_4^{i_4}} \neq 0$$

is $(i_1, i_2, i_3, i_4) = (1, 0, 0, 2)$. To normalize $f_2^{[3]}$, put

$$\begin{aligned}\tilde{x}_i &= a_1 x_i, \quad 1 \leq i \leq 4 \\ \tilde{v} &= a_1 v\end{aligned}$$

where $a_1 = 1/g$. We get the following canonical form for the variable length pendulum

$$\begin{aligned}\dot{\tilde{x}}_1 &= \tilde{x}_2 \\ \dot{\tilde{x}}_2 &= \tilde{x}_3 + \tilde{x}_4^2 \frac{\tilde{x}_1}{1 - \tilde{x}_3^2} \\ \dot{\tilde{x}}_3 &= \tilde{x}_4 \\ \dot{\tilde{x}}_4 &= \tilde{v}\end{aligned}$$

Independently of the value of the gravity constant g , all systems are feedback equivalent to each other.

4.5 Dual Normal Form and Dual m -Invariants

In the normal form $\Sigma_{\text{NF}}^{[m]}$ given by (4.8), all the components of the control vector field $g^{[m-1]}$ are annihilated and all nonremovable nonlinearities are grouped in $f^{[m]}$. Kang and Krener in their pioneering paper [54] have shown that it is possible to transform, via a homogeneous transformation Γ^2 of degree 2, the homogeneous system

$$\Sigma^{[2]} : \dot{\xi} = A\xi + Bu + f^{[2]}(\xi) + g^{[1]}(\xi)u$$

to a dual normal form. In this form, the components of the drift $f^{[2]}$ are annihilated while all nonremovable nonlinearities are, this time, present in $g^{[1]}$. The aim of this section is to propose, for an arbitrary m , a dual normal form for the system $\Sigma^{[m]}$ and a dual normal form for the system Σ^∞ . On the one hand, our dual normal form generalizes, for higher order terms, that given in Ref. [54] for second-order terms, and, on the other hand, dualizes the normal form $\Sigma_{\text{NF}}^{[m]}$. The structure of this section will follow that of Section 4.3: we will present the dual normal form, then we define and study dual m -invariants, and, finally, we give an explicit construction of transformations bringing the system into its dual normal form. For the proofs of all results contained in this section the reader is referred elsewhere [83].

Our first result asserts that we can always bring the homogeneous system $\Sigma^{[m]}$, given by (4.5), into a dual normal form.

THEOREM 10

The homogeneous system $\Sigma^{[m]}$ is equivalent, via a homogeneous feedback transformations Γ^m , to the dual normal form $\Sigma_{\text{DNF}}^{[m]}$ given by

$$\Sigma_{\text{DNF}}^{[m]} : \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 + vx_n Q_{2,n}^{[m-2]}(x_1, \dots, x_n) \\ \vdots \\ \dot{x}_j = x_{j+1} + v \sum_{i=n-j+2}^n x_i Q_{j,i}^{[m-2]}(x_1, \dots, x_i) \\ \vdots \\ \dot{x}_{n-1} = x_n + v \sum_{i=3}^n x_i Q_{j,i}^{[m-2]}(x_1, \dots, x_i) \\ \dot{x}_n = v \end{cases} \quad (4.21)$$

where $Q_{j,i}^{[m-2]}(x_1, \dots, x_i)$ are homogeneous polynomials of degree $m-2$ depending on the indicated variables.

We will give the dual normal form $\Sigma_{\text{DNF}}^{[2]}$ for homogeneous systems of degree two:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 && + vx_n q_{2,n} \\ &\vdots \\ \dot{x}_{n-1} &= x_n + vx_3 q_{n-1,3} + \cdots + vx_n q_{n-1,n} \\ \dot{x}_n &= v\end{aligned}$$

where $q_{j,i} \in \mathbb{R}$.

The following example is a particular case of the previous system and helps illustrate Theorem 10.

Example 8

Consider the system $\Sigma^{[2]}$ defined in $\mathbb{R}^3$ by

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 + \xi_3^2 \\ \dot{\xi}_2 &= \xi_3 \\ \dot{\xi}_3 &= u\end{aligned}$$

It is easy to check that the change of coordinates $x_1 = \xi_1$, $x_2 = \xi_2 + \xi_3^2$, $x_3 = \xi_3$, and $x_4 = \xi_4$ yields the dual normal form ($n = 3$, $q_{2,3} = 2$, and $v = u$)

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 + 2x_3v \\ \dot{x}_3 &= v\end{aligned}$$

Now we will define dual m -invariants. To start with, recall that the homogeneous vector field $X_i^{[m-1]}$ is defined by taking the homogeneous part of degree $m - 1$ of the vector field

$$X_i^{m-1} = (-1)^i \text{ad}_{A\xi + f^{[m]}}^i (B + g^{[m-1]}).$$

By $X_i^{[m-1]}(\pi_i(\xi))$ we will denote $X_i^{[m-1]}$ evaluated at the point $\pi_i(\xi) = (\xi_1, \dots, \xi_i, 0, \dots, 0)^T$ of the subspace

$$W_i = \{\xi = (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n : \xi_{i+1} = \cdots = \xi_n = 0\}.$$

Consider the system $\Sigma^{[m]}$ and for any j , such that $2 \leq j \leq n-1$, define the polynomial $b_j^{[m-1]}$ by setting

$$b_j^{[m-1]} = g_j^{[m-1]} + \sum_{k=1}^{j-1} L_B L_{A\xi}^{j-k-1} f_k^{[m]} - \sum_{i=1}^n L_B L_{A\xi}^{j-1} \int_0^{\xi_i} C X_{n-i}^{[m-1]}(\pi_i(\xi)) d\xi_i$$

The homogeneous polynomials $b_j^{[m-1]}$, for $2 \leq j \leq n-1$, will be called the *dual m -invariants* of the homogeneous system $\Sigma^{[m]}$.

Consider two systems $\Sigma^{[m]}$ and $\tilde{\Sigma}^{[m]}$ of the forms (4.5) and (4.6). Let

$$\left\{ b_j^{[m-1]} : 2 \leq j \leq n-1 \right\}, \quad \text{and} \quad \left\{ \tilde{b}_j^{[m-1]} : 2 \leq j \leq n-1 \right\}$$

denote, respectively, their dual m -invariants. The following result dualizes that of Theorem 4.

THEOREM 11

The dual m -invariants have the following properties:

1. Two systems $\Sigma^{[m]}$ and $\tilde{\Sigma}^{[m]}$ are equivalent via a homogeneous feedback transformation Γ^m if and only if

$$b_j^{[m-1]} = \tilde{b}_j^{[m-1]}, \quad \text{for any } 2 \leq j \leq n-1$$

2. The dual m -invariants $\bar{b}_j^{[m-1]}$ of the dual normal form $\Sigma_{\text{DNF}}^{[m]}$, defined by (4.21), are given by

$$\bar{b}_j^{[m-1]}(x) = \sum_{i=n-j+2}^n x_i Q_{j,i}^{[m-2]}(x_1, \dots, x_i), \quad \text{for any } 2 \leq j \leq n-1$$

This result asserts that the dual m -invariants, similarly to m -invariants, form a set of complete invariants of the homogeneous feedback transformation. Notice, however, that the same information is encoded in both sets of invariants in different ways.

Like Theorem 4 for normal form $\Sigma_{\text{NF}}^{[m]}$, Theorem 11 shows that the polynomial functions $Q_{j,i}^{[m-2]}$ defining the dual normal form $\Sigma_{\text{DNF}}^{[m]}$ are unique

under feedback transformation Γ^m . The question that remains is how to bring a given system into its dual normal form $\Sigma_{\text{DNF}}^{[m]}$. To this end, define the following homogeneous polynomials

$$\begin{aligned}\phi_1^{[m]} &= - \sum_{i=1}^n \int_0^{\xi_i} CX_{n-i}^{[m-1]}(\pi_i(\xi)) \, d\xi_i \\ \phi_{j+1}^{[m]} &= f_j^{[m]} + L_{A\xi} \phi_j^{[m]}, \quad 1 \leq j \leq n-1 \\ \alpha^{[m]} &= - \left(f_n^{[m]} + L_{A\xi} \phi_n^{[m]} \right) \\ \beta^{[m-1]} &= - \left(g_n^{[m-1]} + L_B \phi_n^{[m]} \right)\end{aligned}\tag{4.22}$$

THEOREM 12

The feedback transformation

$$\begin{aligned}\Gamma^m : \quad x &= \xi + \phi^{[m]}(\xi) \\ u &= v + \alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v\end{aligned}$$

where $\alpha^{[m]}$, $\beta^{[m-1]}$, and the components $\phi_j^{[m]}$ of $\phi^{[m]}$ are defined by (4.22), brings the system $\Sigma^{[m]}$ into its dual normal form $\Sigma_{\text{DNF}}^{[m]}$ given by (4.21).

Now our aim is to dualize the normal form $\Sigma_{\text{NF}}^\infty$. Consider the system Σ^∞ of the form (4.16) and assume that its linear part (F, G) is controllable.

Consider the system Σ^∞ of the form (4.2) and recall that we assume the linear part (F, G) to be controllable. Apply a series of transformations Γ^m , $m = 1, 2, 3, \dots$ successively to Σ^∞ , such that each Γ^m brings $\Sigma^{[m]}$ to its dual normal form $\Sigma_{\text{DNF}}^{[m]}$. More precisely, bring (F, G) into the Brunovsky canonical form (A, B) via a linear feedback Γ^1 and denote $\Sigma^{\infty,1} = \Gamma_*^1(\Sigma^\infty)$. Assume that a system $\Sigma^{\infty,m-1}$ has been defined. Let Γ^m be a homogeneous feedback transformation transforming $\Sigma^{[m]}$, which is the homogeneous part of degree m of $\Sigma^{\infty,m-1}$, to the dual normal form $\Sigma_{\text{DNF}}^{[m]}$ (the transformation Γ^m can be taken, for instance, as the transformations defined by (4.22)). Define $\Sigma^{\infty,m} = \Gamma_*^m(\Sigma^{\infty,m-1})$. Notice that we apply Γ^m to the whole system $\Sigma^{\infty,m-1}$ (and not only to its homogeneous part $\Sigma^{[m]}$). Successive iteration of Theorem 12 gives the following dual normal form.

THEOREM 13

The system Σ^∞ can be transformed via a formal feedback transformation Γ^∞ , into the dual normal form $\Sigma_{\text{DNF}}^\infty$ given by

$$\Sigma_{\text{DNF}}^\infty : \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 + vx_n Q_{2,n}(x_1, \dots, x_n) \\ \vdots \\ \dot{x}_j = x_{j+1} + v \sum_{i=n-j+2}^n x_i Q_{j,i}(x_1, \dots, x_i) \\ \vdots \\ \dot{x}_{n-1} = x_n + v \sum_{i=3}^n x_i Q_{j,i}(x_1, \dots, x_i) \\ \dot{x}_n = v \end{cases} \quad (4.23)$$

where $Q_{j,i}(x_1, \dots, x_i)$ are formal power series depending on the indicated variables.

4.6 Dual Canonical Form

Similarly to normal forms, a given system can admit different dual normal forms. We are thus interested in constructing a dual canonical form (which would dualize the canonical form $\Sigma_{\text{CF}}^\infty$ in the same way as $\Sigma_{\text{DNF}}^\infty$ dualizes $\Sigma_{\text{NF}}^\infty$). Assuming that the linear part (F, G) of the system Σ^∞ , of the form (4.16), is controllable, we denote by m_0 the degree of the first homogeneous term of the system Σ^∞ which cannot be annihilated by a feedback transformation. Thus by Theorem 11 and Theorem 12 [using transformations (4.22)], we can assume, after applying a suitable feedback, that Σ^∞ takes the form

$$\Sigma^\infty : \dot{\xi} = A\xi + Bu + \bar{g}^{[m_0-1]}(\xi)u + \sum_{m=m_0+1}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right)$$

where (A, B) is in Brunovský canonical form and the first nonvanishing homogeneous vector field $\bar{g}^{[m_0-1]}$ is in the dual normal form, compare (4.21), with components given by

$$\bar{g}_j^{[m_0-1]}(\xi) = \begin{cases} \sum_{i=n-j+2}^n \xi_i Q_{j,i}^{[m_0-2]}(\xi_1, \dots, \xi_i), & 2 \leq j \leq n-1 \\ 0, & j = 1 \text{ or } j = n \end{cases}$$

Define

$$j_* = \inf \left\{ j = 2, \dots, n-1 : \bar{g}_j^{[m_0-1]}(\xi) \neq 0 \right\}$$

and let $(i_1, \dots, i_n)$, such that $i_1 + \dots + i_n = m_0 - 1$, be the largest n -tuple (in the lexicographic ordering) of nonnegative integers such that

$$\frac{\partial^{m_0-1} \bar{g}_{j_*}^{[m_0-1]}}{\partial \xi_1^{i_1} \dots \partial \xi_n^{i_n}} \neq 0$$

We have the following theorem.

THEOREM 14

There exists a formal feedback transformation Γ^∞ which brings the system Σ^∞ into the following form

$$\Sigma_{\text{DCF}}^\infty : \dot{x} = Ax + Bv + \sum_{m=m_0}^{\infty} \bar{g}^{[m-1]}(x)v$$

where for any $m \geq m_0$, the components $\bar{g}_j^{[m-1]}$ of $\bar{g}^{[m-1]}$ are given by

$$\bar{g}_j^{[m-1]} = \begin{cases} \sum_{i=n-j+2}^n x_i Q_{j,i}^{[m-2]}(x_1, \dots, x_i), & 2 \leq j \leq n-1 \\ 0, & j = 1 \text{ or } j = n \end{cases} \quad (4.24)$$

Moreover,

$$\frac{\partial^{m_0-1} \bar{g}_{j_*}^{[m_0-1]}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}} = \pm 1 \quad (4.25)$$

and for any $m \geq m_0 + 1$

$$\frac{\partial^{m_0-1} \bar{g}_{j_*}^{[m-1]}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}}(x_1, 0, \dots, 0) = 0 \quad (4.26)$$

The form $\Sigma_{\text{DCF}}^\infty$, which satisfies (4.24)–(4.26), will be called dual canonical form of Σ^∞ . The name is justified by the following theorem.

THEOREM 15

Two systems Σ_1^∞ and Σ_2^∞ are formally feedback equivalent if and only if their dual canonical forms $\Sigma_{1,\text{DCF}}^\infty$ and $\Sigma_{2,\text{DCF}}^\infty$ coincide.

Example 9

Consider the system

$$\Sigma : \dot{\xi} = f(\xi) + g(\xi)u, \quad \xi \in \mathbb{R}^3, \quad u \in \mathbb{R}$$

whose linear part is assumed to be controllable. Theorem 13 assures that the system Σ is formally feedback equivalent to the dual normal form $\Sigma_{\text{DNF}}^{\infty}$ given by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 + vx_3Q(x_1, x_2, x_3) \\ \dot{x}_3 &= v \end{aligned}$$

where $Q(x_1, x_2, x_3)$ is a formal power series of the variables x_1, x_2, x_3 .

Assume for simplicity that $m_0 = 2$, which is equivalent to the condition: $g, \text{ad}_f g$, and $[g, \text{ad}_f g]$ linearly independent at $0 \in \mathbb{R}^3$. This implies that we can represent $Q = Q(x_1, x_2, x_3)$, as

$$Q = c + x_1Q_1(x_1) + x_2Q_2(x_1, x_2) + x_3Q_3(x_1, x_2, x_3)$$

where $c \in \mathbb{R}, c \neq 0$.

Observe that any Q of the aforementioned form gives a dual normal form $\Sigma_{\text{DNF}}^{\infty}$. In order to get the dual canonical form we use Theorem 14, which assures that the system Σ is formally feedback equivalent to its dual canonical form $\Sigma_{\text{DCF}}^{\infty}$ defined by

$$\begin{aligned} \dot{\tilde{x}}_1 &= \tilde{x}_2 \\ \dot{\tilde{x}}_2 &= \tilde{x}_3 + \tilde{v}\tilde{x}_3\tilde{Q}(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3) \\ \dot{\tilde{x}}_3 &= \tilde{v} \end{aligned}$$

where $\tilde{Q}(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$ is a formal power series such that

$$\tilde{Q}(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3) = 1 + \tilde{x}_2\tilde{Q}_2(\tilde{x}_1, \tilde{x}_2) + \tilde{x}_3\tilde{Q}_3(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$$

4.7 Normal Forms for Single-Input Systems with Uncontrollable Linearization

4.7.1 Introduction

In Section 4.3 we presented the normal forms $\Sigma_{\text{NF}}^{[m]}$ and $\Sigma_{\text{NF}}^{\infty}$ of the system Σ whose linearization (i.e., linear approximation) is controllable. In this

section, we will deal with systems with uncontrollable linearization. A normal form for homogeneous systems of degree 2, with uncontrollable linearization, was proposed by Kang [51]. The normal form presented in this section was obtained by the authors [81, 85] and it generalizes, on the one hand, the normal form of Kang [51] (uncontrollable linearization) and on the other, the normal form $\Sigma_{\text{NF}}^{[m]}$ (controllable linearization) also obtained by Kang and presented in Section 4.3. Another normal form for single-input systems with uncontrollable linearization has also been proposed by Krener et al. [58–60], and by the authors [78], which differ from ours by another definition of homogeneity (we consider the latter with respect to the linearly controllable variables while theirs is with respect to all variables, see Example 10).

4.7.2 Taylor Series Expansions

All objects (i.e., functions, maps, vector fields, control systems, etc.) are considered in a neighborhood of $0 \in \mathbb{R}^n$ and assumed to be C^∞ -smooth. Consider the single-input system

$$\Sigma : \dot{\xi} = f(\xi) + g(\xi)u, \quad \xi \in \mathbb{R}^n, u \in \mathbb{R}$$

We assume throughout this section that $f(0) = 0$ and $g(0) \neq 0$. Let

$$\Sigma^{[1]} : \dot{\xi} = F\xi + Gu$$

where $F = (\partial f / \partial \xi)(0)$ and $G = g(0)$, be the linear approximation of the system around the equilibrium point $0 \in \mathbb{R}^n$. If the linear approximation is not controllable, which is the case studied in this section, then there exists a positive integer $r \in \mathbb{N}$ such that

$$\text{rank}(G, FG, \dots, F^{n-1}G) = n - r$$

Moreover, there exist coordinates $\xi = (\xi_1, \dots, \xi_r, \xi_{r+1}, \dots, \xi_n)^T$ of $\mathbb{R}^r \times \mathbb{R}^{n-r}$ in which the pair (F, G) admits the following Kalman decomposition

$$A = \begin{pmatrix} F_1 & 0 \\ F_3 & F_2 \end{pmatrix}_{n \times n} \quad \text{and} \quad B = \begin{pmatrix} 0 \\ G_2 \end{pmatrix}_{n \times 1}$$

where the pair (F_2, G_2) is controllable. Throughout this section, r will stand for the dimension of the uncontrollable part of the linear approximation of the system and $\xi = (\xi_1, \dots, \xi_r, \xi_{r+1}, \dots, \xi_n)^T$ will denote coordinates defining the Kalman decomposition.

We will use the notation $C_0^\infty(\mathbb{R}^r)$ for the space of germs at $0 \in \mathbb{R}^r$ of smooth $\mathbb{R}$ -valued functions of $\xi = (\xi_1, \dots, \xi_r)^T \in \mathbb{R}^r$. By $\mathbb{R}[[\xi_1, \dots, \xi_r]]$, we will denote the space of formal power series of $\xi_1, \dots, \xi_r$ with coefficients in $\mathbb{R}$.

Let h be a smooth $\mathbb{R}$ -valued function defined in a neighborhood of $0 \times 0 \in \mathbb{R}^r \times \mathbb{R}^{n-r}$. By

$$h(\xi) = h^{[0]}(\xi) + h^{[1]}(\xi) + h^{[2]}(\xi) + \dots = \sum_{m=0}^{\infty} h^{[m]}(\xi)$$

we denote its Taylor series expansion with respect to $(\xi_{r+1}, \dots, \xi_n)^T$ at $0 \in \mathbb{R}^r \times \mathbb{R}^{n-r}$, where $h^{[m]}(\xi)$ stands for a homogeneous polynomial of degree m of the variables $\xi_{r+1}, \dots, \xi_n$ whose coefficients are in $C_0^\infty(\mathbb{R}^r)$.

Similarly, throughout this section, for a map ϕ of an open subset of $\mathbb{R}^r \times \mathbb{R}^{n-r}$ to $\mathbb{R}^r \times \mathbb{R}^{n-r}$ (resp. for a vector field f on an open subset of $\mathbb{R}^r \times \mathbb{R}^{n-r}$) we will denote by $\phi^{[m]}$ (resp. $f^{[m]}$) the term of degree m of its Taylor series expansion with respect to $(\xi_{r+1}, \dots, \xi_n)^T$ at $0 \in \mathbb{R}^r \times \mathbb{R}^{n-r}$, that is, each component $\phi_j^{[m]}$ of $\phi^{[m]}$ (resp. $f_j^{[m]}$ of $f^{[m]}$) is a homogeneous polynomial of degree m of the variables $\xi_{r+1}, \dots, \xi_n$ whose coefficients are in $C_0^\infty(\mathbb{R}^r)$.

Consider the Taylor series expansion of the system Σ given by

$$\Sigma^\infty : \dot{\xi} = F\xi + Gu + f^{[0]}(\xi) + \sum_{m=1}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}u \right) \quad (4.27)$$

Note that, although we assume $f(0) = 0$, the term $f^{[0]}(\xi)$ is present because the degree is computed with respect to the variables $\xi_{r+1}, \dots, \xi_n$ only and thus $f^{[0]}$ is, in general, a function of $\xi_1, \dots, \xi_r$.

Consider also the Taylor series expansion Γ^∞ of the feedback transformation Γ given by

$$\begin{aligned} x &= T\xi + \sum_{m=0}^{\infty} \phi^{[m]}(\xi) \\ \Gamma^\infty : \\ u &= K\xi + Lv + \alpha^{[0]}(\xi) + \sum_{m=1}^{\infty} \left(\alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v \right) \end{aligned} \quad (4.28)$$

where T is an invertible matrix and $L \neq 0$.

The method proposed by Kang and Krener is to study the action of Γ^∞ on the system Σ^∞ step by step, that is, to analyze successively the action of the homogeneous parts of Γ^∞ on the homogeneous parts, of the same degree,

of Σ^∞ . Notice, however, that in their approach, Kang and Krener consider Taylor series expansions with respect to all state variables $\xi_1, \dots, \xi_n$ of the system and, as a consequence, homogeneity is considered with respect to all variables $\xi_1, \dots, \xi_n$. Following our approach [81, 85] we propose a slight modification of this homogeneity. In view of a different nature of the controllable and uncontrollable parts of the linear approximation, we consider Taylor series expansions with respect to the linearly controllable variables $\xi_{r+1}, \dots, \xi_n$ only. This leads to considering as homogeneous parts of the system and of the feedback transformations, according to our definition, terms that are polynomial with respect to $\xi_{r+1}, \dots, \xi_n$ with smooth coefficients depending on $\xi_1, \dots, \xi_r$. When analyzing the action of a homogeneous transformation Γ^m (understood as homogeneity with respect to the controllable variables) on the system Σ^∞ , we can notice three undesirable phenomena (see Section 4.7.5) that are not present in the action of homogeneous transformations in the controllable case (where homogeneity is considered with respect to all variables). To deal with these problems caused by the presence of the uncontrollable linear part, we will introduce, in Section 4.7.5, different weights for the components corresponding to the controllable and uncontrollable parts.

4.7.3 Linear Part and Resonances

Let $\lambda = (\lambda_1, \dots, \lambda_r) \in \mathbb{C}^r$ be the set of eigenvalues associated to the uncontrollable part of the linear system

$$\dot{\xi} = F\xi + Gu$$

By a linear feedback transformation

$$\Gamma^1 : \begin{aligned} x &= T\xi \\ u &= K\xi + Lv \end{aligned}$$

it is always possible to bring the linear system into the following Jordan–Brunovský canonical form

$$A = \begin{pmatrix} J & 0 \\ 0 & A_2 \end{pmatrix}_{n \times n} \quad \text{and} \quad B = \begin{pmatrix} 0 \\ B_2 \end{pmatrix}_{n \times 1}$$

where J is the Jordan canonical form of dimension r and (A_2, B_2) the Brunovský canonical form of dimension $n - r$. In the case when all

eigenvalues λ_i are real, we have

$$J = \begin{pmatrix} \lambda_1 & \sigma_2 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_r \\ 0 & \cdots & 0 & \lambda_r \end{pmatrix}_{r \times r}, \quad \sigma_i \in \{0, 1\}$$

In the case of complex eigenvalues, we replace in J the eigenvalue λ_j by the matrix

$$\Lambda_j = \begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix}_{2 \times 2}$$

where $\lambda_j = \alpha_j + i\beta_j$, and we replace the integer $\sigma_j \in \{0, 1\}$ either by the matrix

$$\Xi_j = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}_{2 \times 2} \quad \text{or} \quad \Xi_j = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{2 \times 2}$$

Recall from Section 4.2 the notion of a resonant eigenvalue and the resonant set associated with it.

DEFINITION 1 An eigenvalue λ_j is called *resonant* if there exists a r -tuple $k = (k_1, \dots, k_r) \in \mathbb{N}^r$ of nonnegative integers, satisfying $|k| = k_1 + \dots + k_r \geq 2$, such that

$$\lambda_j = \lambda_1 k_1 + \dots + \lambda_r k_r$$

For each eigenvalue λ_j , where $1 \leq j \leq r$, we define

$$\mathcal{R}_j = \{k = (k_1, \dots, k_r) \in \mathbb{N}^r : |k| \geq 2 \text{ and } \lambda_j = \lambda_1 k_1 + \dots + \lambda_r k_r\}$$

which is called the *resonant set* associated to λ_j .

4.7.4 Notations and Definitions

The method described in Section 4.3 (proposed by Kang and Krener [54], and then followed by Kang [50, 51] and by the authors [79, 83]) is to analyze step by step the action of the transformation Γ^∞ on the system Σ^∞ . In the controllable case, it consists of bringing the linear part of the system into the

Brunovský canonical form and then applying, step by step, homogeneous feedback transformations of the form

$$\begin{aligned} x &= \xi + \phi^{[m]}(\xi) \\ \Gamma^m : \\ u &= v + \alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v \end{aligned}$$

in order to normalize the homogeneous part of degree m of the system. The advantage of this method, in the controllable case, follows from the fact that homogeneous transformations Γ^m leave invariant all terms of degree smaller than m of the system. The situation gets very different in the uncontrollable case with the modified notion of homogeneity. To see the difference, we analyze the action of Γ^m on the system Σ^∞ , given by (4.27). Let

$$\tilde{\Sigma}^\infty : \dot{x} = Fx + Gv + \tilde{f}^{[0]}(x) + \sum_{m=1}^{\infty} \left(\tilde{f}^{[m]}(x) + \tilde{g}^{[m-1]}(x)v \right)$$

be the system Σ^∞ transformed by Γ^m . Recall that for both systems, the first r components of the state correspond to the uncontrollable part and that the degree of homogeneity (for all terms of the systems and for the transformation Γ^m) is computed with respect to the controllable variables, that is, the last $n - r$ variables only. We can observe the following three undesirable phenomena.

First, note that the homogeneous transformation Γ^m does not preserve homogeneous (with respect to linearly controllable variables) terms of degree smaller than m . It only preserves terms of degree smaller than $m - 1$, that is, $\tilde{f}^{[k]} = f^{[k]}$ and $\tilde{g}^{[k-1]} = g^{[k-1]}$, for any $0 \leq k \leq m - 2$, while it transforms those of degree $m - 1$ as follows

$$\tilde{g}^{[m-2]} = g^{[m-2]} \quad \text{and} \quad \tilde{f}^{[m-1]} = f^{[m-1]} + \sum_{i=r+1}^n \frac{\partial \phi^{[m]}}{\partial \xi_i} f_i^{[0]}$$

Note that, when comparing the homogeneous parts of the same degree k of two systems, we have to compare the homogeneous parts of degree k of the drifts and the homogeneous parts of degree $k - 1$ of control vector fields. As the homogeneity is considered with respect to the state and the control, the homogeneous part of degree k of the system is represented by the homogeneous part $f^{[k]}$, of degree k , of the drift and the homogeneous part $g^{[k-1]}$, of degree $k - 1$, of the control vector field multiplied by the control u .

Secondly, Γ^m transforms the homogeneous part $(f^{[m]}, g^{[m-1]})$, of degree m , not according to a homological equation, but in such a way that

$$\begin{aligned}\tilde{f}^{[m]} &= f^{[m]} + [A\xi, \phi^{[m]}] + B\alpha^{[m]} + \sum_{i=r+1}^n \left(\frac{\partial \phi^{[m]}}{\partial \xi_i} f_i^{[1]} - \frac{\partial f^{[1]}}{\partial \xi_i} \phi_i^{[m]} \right) \\ &\quad + \sum_{i=1}^r \left(\frac{\partial \phi^{[m]}}{\partial \xi_i} f_i^{[0]} - \frac{\partial f^{[0]}}{\partial \xi_i} \phi_i^{[m]} \right) + g^{[0]} \alpha^{[m]} \\ \tilde{g}^{[m-1]} &= g^{[m-1]} + \text{ad}_B \phi^{[m]} + B\beta^{[m-1]} + g^{[0]} \beta^{[m-1]} + \sum_{i=r+1}^n \frac{\partial \phi^{[m]}}{\partial \xi_i} g_i^{[0]}\end{aligned}$$

Thirdly, the Lie bracket of two homogeneous vectors fields $f^{[m]}$ and $g^{[k]}$ is given by

$$[f^{[m]}, g^{[k]}](\xi) = \sum_{i=1}^n \left(\frac{\partial g^{[k]}}{\partial \xi_i} f_i^{[m]}(\xi) - \frac{\partial f^{[m]}}{\partial \xi_i} g_i^{[k]}(\xi) \right)$$

and thus fails, in general, to be homogeneous of degree $m + k - 1$ because the terms $(\partial g^{[k]} / \partial \xi_i) f_i^{[m]}(\xi)$ and $(\partial f^{[m]} / \partial \xi_i) g_i^{[k]}(\xi)$, for $1 \leq i \leq r$, are, in general, homogeneous of degree $m + k$.

These three inconveniences are caused only by the fact that differentiating with respect to the variables $\xi_1, \dots, \xi_r$ does not decrease the degree (in particular, by the presence of terms of degree 0 with respect to the variables $\xi_{r+1}, \dots, \xi_n$). To overcome this, we define, for any $m \geq 0$,

$$\begin{aligned}f^{(m)} &= (f_1^{[m-1]}, \dots, f_r^{[m-1]}, f_{r+1}^{[m]}, \dots, f_n^{[m]})^T \\ g^{(m)} &= (g_1^{[m-1]}, \dots, g_r^{[m-1]}, g_{r+1}^{[m]}, \dots, g_n^{[m]})^T \\ \phi^{(m)} &= (\phi_1^{[m-1]}, \dots, \phi_r^{[m-1]}, \phi_{r+1}^{[m]}, \dots, \phi_n^{[m]})^T\end{aligned}$$

where, for any $1 \leq i \leq r$, we set $f_i^{[-1]} = g_i^{[-1]} = \phi_i^{[-1]} = 0$.

Control systems, vector fields, feedback transformations, etc., that are homogeneous with respect to the just-defined weights, will be called *weighted homogeneous*. Note that the weighted homogeneity just defined is related with the decomposition of the state space $\mathbb{R}^n$ into uncontrollable and controllable parts and therefore it does not apply to applications with values in $\mathbb{R}$. In particular, for real-valued homogeneous polynomials $h^{[m]}$, we will write $h^{(m)} = h^{[m]}$.

We will denote by $P^{[m]}(\xi)$ the space of homogeneous polynomials of degree m of the variables $\xi_{r+1}, \dots, \xi_n$ (with coefficients depending on $\xi_1, \dots, \xi_r$) and by $P^{\geq m}(\xi)$ the space of formal power series of the variables $\xi_{r+1}, \dots, \xi_n$ (with coefficients depending on $\xi_1, \dots, \xi_r$) starting from terms of degree m .

We will denote by $V^{(m)}(\xi)$ the space of weighted $\langle m \rangle$ -homogeneous vector fields, that is, the space of vector fields whose first r components are in $P^{[m-1]}(\xi)$ and the last $n - r$ components are in $P^{[m]}(\xi)$. Moreover, $V^{\langle \geq m \rangle}(\xi)$ will denote the space of vector fields formal power series whose first r components are in $P^{\geq m-1}(\xi)$ and the last $n - r$ components are in $P^{\geq m}(\xi)$.

4.7.5 Weighted Homogeneous Systems

Applying a linear feedback transformation, we can bring the linear approximation (F, G) of the system into the Jordan–Brunovský canonical form (A, B) , that is, the uncontrollable part, of dimension r , is in the Jordan form and the controllable part, of dimension $n - r$, in the Brunovský form. Notice, however, that contrary to the controllable case (where there are no zero degree terms while the terms of degree one are just linear terms that we bring to the Brunovský canonical form), after having normalized linear terms of an uncontrollable system, we are still left with weighted homogeneous terms of degree zero and one. We can normalize them as follows.

PROPOSITION 4

Consider the system

$$\Sigma^{\langle \leq 1 \rangle} : \dot{\xi} = A\xi + Bu + f^{(0)}(\xi) + f^{(1)}(\xi) + g^{(0)}(\xi)u$$

where (A, B) is in the Jordan–Brunovský canonical form:

1. There exists a smooth feedback transformation of the form

$$\begin{aligned} \Gamma^{\langle \leq 1 \rangle} : \quad & x = \xi + \phi^{(0)}(\xi) + \phi^{(1)}(\xi) \\ & u = v + \alpha^{(0)}(\xi) + \alpha^{(1)}(\xi) + \beta^{(0)}(\xi)v \end{aligned} \quad (4.29)$$

which takes the system $\Sigma^{\langle \leq 1 \rangle}$ into the system

$$\tilde{\Sigma}^{\langle \leq 1 \rangle} : \dot{x} = Ax + Bv + \tilde{f}^{(1)}(x) \quad (4.30)$$

modulo terms in $V^{\langle \geq 2 \rangle}$, where the vector field $\tilde{f}^{(1)}$ satisfies $\tilde{f}_j^{(1)} = 0$ for $r + 1 \leq j \leq n$.

2. Assume that all eigenvalues $\lambda_1, \dots, \lambda_r$ of the Jordan–Brunovský canonical form (A, B) are real and distinct (in particular all $\sigma_j = 0$). Then a formal

feedback transformation $\Gamma^{(1)}$ of the form (4.29), which takes the system $\Sigma^{(\leq 1)}$ into the system $\tilde{\Sigma}^{(\leq 1)}$ defined by (4.30), can be chosen in such a way that the vector field $\tilde{f}^{(1)}$ is replaced by $\tilde{f}^{(1)}(x)$ whose components are given by

$$\tilde{f}_j^{(1)}(x) = \begin{cases} \sum_{k \in \mathcal{R}_j} \gamma_j^k x_1^{k_1} \cdots x_r^{k_r} & \text{if } 1 \leq j \leq r \\ 0 & \text{if } r+1 \leq j \leq n \end{cases} \quad (4.31)$$

The existence of a transformation $x = \xi + \phi^{(1)}(\xi)$ yielding the form of $\tilde{f}_j^{(1)}(x)$ in (2) is an immediate consequence of Theorem 1.

Now we will study the action of the *weighted homogeneous feedback*

$$\Gamma^{(m)} : \begin{cases} x = \xi + \phi^{(m)}(\xi) \\ u = v + \alpha^{(m)}(\xi) + \beta^{(m-1)}(\xi)v \end{cases} \quad (4.32)$$

on the following *weighted homogeneous system*

$$\Sigma^{(m)} : \dot{\xi} = A\xi + Bu + \tilde{f}^{(1)}(\xi) + f^{(m)}(\xi) + g^{(m-1)}(\xi)u \quad (4.33)$$

where, because of Proposition 4, we assume that $\tilde{f}_j^{(1)}$ is of the form (4.31). Note that a weighted homogeneous feedback is a smooth feedback; it depends polynomially on the variables $\xi_{r+1}, \dots, \xi_n$ and smoothly on the variables $\xi_1, \dots, \xi_r$.

Consider another weighted homogeneous system $\tilde{\Sigma}^{(m)}$ given by

$$\tilde{\Sigma}^{(m)} : \dot{x} = Ax + Bv + \tilde{f}^{(1)}(x) + \tilde{f}^{(m)}(x) + \tilde{g}^{(m-1)}(x)v \quad (4.34)$$

where $\tilde{f}^{(1)} = \tilde{f}^{(1)}$. We will say that $\Gamma^{(m)}$ transforms $\Sigma^{(m)}$ into $\tilde{\Sigma}^{(m)}$, and we will denote it by $\Gamma_*^{(m)}(\Sigma^{(m)}) = \tilde{\Sigma}^{(m)}$, if $\Gamma^{(m)}$ transforms $\Sigma^{(m)}$ into

$$\dot{x} = Ax + Bv + \tilde{f}^{(1)}(x) + \tilde{f}^{(m)}(x) + \tilde{g}^{(m-1)}(x)v + R^{(\geq m+1)}(x, v)$$

where $R^{(\geq m+1)}(x, v) \in V^{(\geq m+1)}(x, v)$.

Recall that λ_j , for $1 \leq j \leq r$, are the eigenvalues of the uncontrollable part of the linear approximation and that σ_j for $2 \leq j \leq r$, define the corresponding Jordan form (see Section 4.7.3). We define additionally $\sigma_{r+1} = 0$ and for any $r+1 \leq j \leq n-1$, we put $\lambda_j = 0$ and $\sigma_{j+1} = 1$.

Analysis of weighted homogeneous systems is based on the following result which generalizes, to the uncontrollable case, that proved by Kang [50] (and recalled in Proposition 2).

PROPOSITION 5

For any $m \geq 2$, the feedback transformation $\Gamma^{(m)}$, of the form (4.32), brings the system $\Sigma^{(m)}$, given by (4.33), into $\tilde{\Sigma}^{(m)}$, given by (4.34), if and only if the following relations

$$\left\{ \begin{array}{l} L_{A\xi+f^{(1)}}\phi_j^{(m)} - \lambda_j\phi_j^{(m)} - \sigma_{j+1}\phi_{j+1}^{(m)} = \tilde{f}_j^{(m)} - f_j^{(m)} \\ L_B\phi_j^{(m)} = \tilde{g}_j^{(m-1)} - g_j^{(m-1)} \\ L_{A\xi+f^{(1)}}\phi_n^{(m)} + \alpha^{(m)} = \tilde{f}_n^{(m)} - f_n^{(m)} \\ L_B\phi_n^{(m)} + \beta^{(m-1)} = \tilde{g}_n^{(m-1)} - g_n^{(m-1)} \end{array} \right. \quad (4.35)$$

hold for any $1 \leq j \leq n-1$.

This proposition can be viewed as the *weighted control homological equation* for systems with uncontrollable linearization (its proof follows the same line as that of Kang [50] for the standard control homological equation (CHE)). Once again, solving a system of first-order partial differential equations may be avoided if the analysis is performed step by step, and thus the feedback equivalence of two systems $\Sigma^{(m)}$ and $\tilde{\Sigma}^{(m)}$ (with uncontrollable linearization) is reduced to solving the algebraic system (4.35).

The following result gives our normal form for weighted homogeneous systems with uncontrollable linearization. Recall the notation $\pi_i(x) = (x_1, \dots, x_i)$.

THEOREM 16

For any $m \geq 2$, there exists a weighted feedback transformation $\Gamma^{(m)}$ that transforms the weighted homogeneous system $\Sigma^{(m)}$, given by (4.33), into its weighted homogeneous normal form

$$\Sigma_{\text{NF}}^{(m)} : \dot{x} = Ax + Bv + \bar{f}^{(1)}(x) + \bar{f}^{(m)}(x) \quad (4.36)$$

where $\bar{f}^{(1)}(x)$ is given by Proposition 4 (2) and the vector field $\bar{f}^{(m)}(x)$ satisfies

$$\bar{f}_j^{(m)}(x) = \begin{cases} x_{r+1}^{m-1} S_{j,m}(\pi_r(x)) + \sum_{i=r+2}^n x_i^2 Q_{j,i}^{(m-3)}(\pi_i(x)) & \text{if } 1 \leq j \leq r \\ \sum_{i=j+2}^n x_i^2 P_{j,i}^{(m-2)}(\pi_i(x)) & \text{if } r+1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \quad (4.37)$$

where $S_{j,m}(\pi_r(x)) \in C_0^\infty(\mathbb{R}^r)$ are C^∞ -functions of the variables $x_1, \dots, x_r$ and the functions $P_{j,i}^{(m-2)}$ and $Q_{j,i}^{(m-3)}$ are homogeneous polynomials, respectively, of degree $m-2$ and $m-3$, of the variables $x_{r+1}, \dots, x_i$, with coefficients in $C_0^\infty(\mathbb{R}^r)$.

The proof of Theorem 16 is based on Theorem 18, stated in Section 4.7.7, which explicitly gives transformations bringing $\Sigma^{(m)}$ into its normal form $\Sigma_{\text{NF}}^{(m)}$. The normal form generalizes to the uncontrollable case the normal form $\Sigma_{\text{NF}}^{[m]}$ of Kang [50] for systems with controllable linearization, which was stated in Theorem 3. It can also be viewed as a generalization of the normal form obtained by Kang [51] in the uncontrollable case for second-order terms. Other normal forms, for systems with uncontrollable linearization, have been obtained earlier [78], and for third-order terms also [58–60]. Note, however, that those normal forms coincide neither with our normal form $\Sigma_{\text{NF}}^{(2)}$ nor $\Sigma_{\text{NF}}^{(3)}$ because of the different weights used, as explained in the following example.

Example 10

Consider the system $\dot{\xi} = f(\xi) + g(\xi)u$ on $\mathbb{R}^3$ and assume that the linearly controllable subsystem is two-dimensional and that the linear part is in the Jordan–Brunovský canonical form. The homogenous system $\Sigma^{[2]}$ (homogeneity being calculated with respect to all variables ξ_1, ξ_2, ξ_3) is

$$\begin{aligned}\dot{\xi}_1 &= \lambda \xi_1 + f_1^{[2]}(\xi) + g_1^{[1]}(\xi)u \\ \dot{\xi}_2 &= \xi_3 + f_2^{[2]}(\xi) + g_2^{[1]}(\xi)u \\ \dot{\xi}_3 &= u + f_3^{[2]}(\xi) + g_3^{[1]}(\xi)u\end{aligned}$$

The linearly uncontrollable subsystem is of dimension one (with ξ_1 being the linearly uncontrollable variable and $(\xi_2, \xi_3)^T$ being the linearly controllable variables) and the resonant set associated with the eigenvalue λ is empty if and only if $\lambda \neq 0$. Kang [51] proved that $\Sigma^{[2]}$ is equivalent via a homogeneous feedback Γ^2 to the following normal form $\Sigma_{\text{NF}}^{[2]}$:

$$\begin{aligned}\dot{x}_1 &= \lambda x_1 + \gamma_2 x_1^2 + x_2 s_{1,2}(x_1) + x_3^2 q_{1,3} \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u\end{aligned}$$

where $\gamma_2 = 0$ if $\lambda \neq 0$ (no resonances) and $\gamma_2 \in \mathbb{R}$ if $\lambda = 0$. Moreover, $s_{1,2}$ is a linear function of x_1 and $q_{1,3}$ is a constant.

Now we will compare this normal form with the normal forms $\Sigma_{\text{NF}}^{(2)}$ and $\Sigma_{\text{NF}}^{(3)}$. To this end, we start with

$$\begin{aligned}\dot{\xi}_1 &= \lambda \xi_1 + f_1^{(1)}(\xi) \\ \Sigma^{(1)} : \dot{\xi}_2 &= \xi_3 + f_2^{(1)}(\xi) \\ \dot{\xi}_3 &= u + f_3^{(1)}(\xi)\end{aligned}$$

where $f_2^{(1)}(\xi)$ and $f_3^{(1)}(\xi)$ are linear functions with respect to ξ_2 and ξ_3 with coefficients that are arbitrary functions of ξ_1 while $f_1^{(1)}(\xi)$ is an arbitrary function of ξ_1 whose development at zero starts with quadratic terms. Now Proposition 4(1) implies that we can annihilate $f_2^{(1)}$ and $f_3^{(1)}$. If $\lambda \neq 0$, then there are no resonances so we can also annihilate $f_1^{(1)}$ and without loss of generality we can assume that $f_1^{(1)}(\xi)$ is in the normal form $\bar{f}_1^{(1)}(\xi) = 0$ assured by Proposition 4. If $\lambda = 0$, then all terms of $f_1^{(1)}$ are resonant so we can assume that $f_1^{(1)}$ is in the normal form $\bar{f}_1^{(1)}(\xi) = (\sum_{i=2}^{\infty} \gamma_i \xi_1^i)$, where $\gamma_i \in \mathbb{R}$. The vector field $(\sum_{i=2}^{\infty} \gamma_i \xi_1^i)(\partial/\partial \xi_1)$ can be normalized by a local diffeomorphism around $\xi_1 = 0$ into the form $(\pm x_1^p + \gamma_{2p-1} x_1^{2p-1})(\partial/\partial x_1)$; however, this will not be used here.

Consider the system

$$\Sigma^{(2)} : \dot{\xi} = A\xi + Bv + \bar{f}^{(1)}(\xi) + f^{(2)}(\xi) + g^{(1)}(\xi)v$$

where (A, B) is in the Jordan–Brunovský canonical form and $\bar{f}^{(1)}$ is in the aforementioned normal form, that is $A\xi + \bar{f}^{(1)}(\xi) + Bu = (\lambda \xi_1 + \bar{f}_1^{(1)}(\xi))(\partial/\partial \xi_1) + \xi_3(\partial/\partial \xi_2) + u(\partial/\partial \xi_3)$, where $f_1^{(1)}(\xi)$ equals 0 or $\sum_{i=2}^{\infty} \gamma_i \xi_1^i$ (depending on λ). Moreover, the components $f_2^{(2)}(\xi)$ and $f_3^{(2)}(\xi)$ are quadratic functions of ξ_2 and ξ_3 ; the components $f_1^{(2)}(\xi)$, $g_2^{(1)}(\xi)$, and $g_3^{(1)}(\xi)$ are linear functions of ξ_2, ξ_3 (all coefficients depending on ξ_1), and $g_1^{(1)}(\xi)$ depends only on ξ_1 . By a weighted homogenous (of degree 2 with respect to ξ_2, ξ_3) feedback transformation $\Gamma^{(2)}$, we can bring $\Sigma^{(2)}$ into the normal form

$$\begin{aligned} \dot{x}_1 &= \lambda x_1 + \bar{f}_1^{(1)}(x) + x_2 S_{1,2}(x_1) \\ \Sigma_{\text{NF}}^{(2)} : \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u \end{aligned}$$

where $S_{1,2}$ is an arbitrary function of x_1 .

Now consider

$$\Sigma^{(3)} : \dot{\xi} = A\xi + Bu + \bar{f}^{(1)}(\xi) + f^{(3)}(\xi) + g^{(2)}(\xi)u$$

where (A, B) is in the Jordan–Brunovský canonical form and $\bar{f}^{(1)}$ is in the normal form described earlier. Moreover, the components $f_2^{(3)}(\xi)$ and $f_3^{(3)}(\xi)$ are cubic functions of ξ_2, ξ_3 ; the components $f_1^{(3)}(\xi)$, $g_2^{(2)}(\xi)$, and $g_3^{(2)}(\xi)$ are quadratic functions of ξ_2, ξ_3 ; and $g_1^{(2)}(\xi)$ is a linear function of ξ_2, ξ_3 (all coefficients depending on ξ_1). By a weighted homogenous (of degree 3

with respect to ξ_2, ξ_3) feedback transformation $\Gamma^{(3)}$ we can bring $\Sigma^{(3)}$ into the normal form

$$\begin{aligned} \dot{x}_1 &= \lambda x_1 + \bar{f}_1^{(1)}(x) + x_2^2 S_{1,3}(x_1) + x_3^2 Q_{1,3}(x_1) \\ \Sigma_{\text{NF}}^{(3)} : \dot{x}_2 &= x_3 \\ \dot{x}_3 &= v \end{aligned}$$

where $S_{1,3}$ and $Q_{1,3}$ are arbitrary functions of x_1 .

Observe that the term $x_2 S_{1,2}(x_1)$ of the Kang normal form $\Sigma_{\text{NF}}^{[2]}$, where $S_{1,2}(x_1)$ is linear, shows up as the first term in the development of $x_2 S_{1,2}(x_1)$ of $\Sigma_{\text{NF}}^{(2)}$ while the term $x_3^2 q_{1,3}(x_1)$, where $q_{1,3}$ is constant, shows up as the first term in the development of $x_3^2 Q_{1,3}(x_1)$ but both $S_{1,2}(x_1)$ and $Q_{1,3}(x_1)$ contain, in general, terms of arbitrary degrees (except for constant terms in $S_{1,2}$). This example illustrates mutual differences between the form $\Sigma_{\text{NF}}^{[2]}$ of Kang and of ours $\Sigma_{\text{NF}}^{(2)}$ and $\Sigma_{\text{NF}}^{(3)}$.

4.7.6 Weighted Homogeneous Invariants

In this section, we will define weighted homogeneous invariants $a^{(m)j,i+2}$ of weighted homogeneous systems $\Sigma^{(m)}$ under weighted homogeneous feedback transformations $\Gamma^{(m)}$ and we will state for them results of [81, 85] generalizing, to the uncontrollable case, a result established by Kang [50] in the controllable case.

Consider the weighted homogeneous system (4.33). For any $i \geq 0$, let us define the vector field

$$X_i^{m-1}(\xi) = (-1)^i \text{ad}_{A\xi+f^{(1)}(\xi)+f^{(m)}(\xi)}^i (B + g^{(m-1)}(\xi))$$

and let $X_i^{(m-1)}$ be the homogeneous part of degree $\langle m-1 \rangle$ of X_i^{m-1} . It means that the first r components are homogeneous of degree $m-2$ and the last $n-r$ components are homogeneous of degree $m-1$ with respect to the variables $\xi_{r+1}, \dots, \xi_n$. One can easily check that

$$X_i^{(m-1)}(\xi) = (-1)^i \left(\text{ad}_{A\xi+f^{(1)}(\xi)}^i g^{(m-1)} + \sum_{k=1}^i (-1)^k \text{ad}_{A\xi+f^{(1)}(\xi)}^{i-k} \text{ad}_{A^{k-1}B} f^{(m)} \right)$$

Define the set of indices $\Delta_r = \Delta_r^1 \cup \Delta_r^2 \subset \mathbb{N} \times \mathbb{N}$ by taking

$$\begin{aligned} \Delta_r^1 &= \{(j, i) \in \mathbb{N} \times \mathbb{N} : 1 \leq j \leq r \text{ and } 0 \leq i \leq n-r-1\}, \\ \Delta_r^2 &= \{(j, i) \in \mathbb{N} \times \mathbb{N} : r+1 \leq j \leq n-2 \text{ and } 0 \leq i \leq n-j-2\} \end{aligned}$$

For any $r + 1 \leq k \leq n$, define the following subspaces

$$W_k = \left\{ (\xi_1, \dots, \xi_r, \xi_{r+1}, \dots, \xi_n)^T \in \mathbb{R}^r \times \mathbb{R}^{n-r} : \xi_{k+1} = \dots = \xi_n = 0 \right\}$$

and let $\pi_k(\xi)$ denote the projection on W_k , that is, $\pi_k(\xi) = (\xi_1, \dots, \xi_r, \xi_{r+1}, \dots, \xi_k, 0, \dots, 0)^T$.

For any $1 \leq j \leq n$, we denote by

$$C_j = (0, \dots, 0, 1, 0, \dots, 0)$$

the row vector in $\mathbb{R}^n$, all of whose components are zero except the j th component which equals 1.

For any $(j, i) \in \Delta_r^1$ (resp. $(j, i) \in \Delta_r^2$), we define $a^{(m)j,i+2}(\xi)$ as the homogeneous part of degree $\langle m - 3 \rangle$ (resp. of degree $\langle m - 2 \rangle$) of the function

$$C_j \left[X_i^{m-1}, X_{i+1}^{m-1} \right] (\pi_{n-i}(\xi))$$

One can easily establish that

$$a^{(m)j,i+2}(\xi) = C_j \left(\text{ad}_{A^i B} X_{i+1}^{(m-1)} - \text{ad}_{A^{i+1} B} X_i^{(m-1)} \right) (\pi_{n-i}(\xi))$$

The functions $a^{(m)j,i+2}$ thus defined will be called *weighted homogeneous $\langle m \rangle$ -invariants* of the weighted homogeneous system $\Sigma^{(m)}$.

Consider, along with $\Sigma^{(m)}$ defined by (4.33), the system $\tilde{\Sigma}^{(m)}$ given by (4.34) and denote by $\tilde{a}^{(m)j,i+2}$ the weighted homogeneous $\langle m \rangle$ -invariants of the latter system.

The following result, which generalizes that obtained by Kang [50] for systems with controllable linearization, asserts that the weighted homogeneous $\langle m \rangle$ -invariants $a^{(m)j,i+2}$ are complete invariants of weighted homogeneous feedback and also illustrates their meaning for the normal form $\Sigma_{\text{NF}}^{(m)}$.

THEOREM 17

For any $m \geq 2$, we have the following properties:

1. Two weighted homogeneous systems $\Sigma^{(m)}$, given by (4.33), and $\tilde{\Sigma}^{(m)}$, given by (4.34), are equivalent via a weighted homogeneous feedback $\Gamma^{(m)}$ if and only if, for any $(j, i) \in \Delta_r$, we have

$$a^{(m)j,i+2} = \tilde{a}^{(m)j,i+2}$$

2. The $\langle m \rangle$ -invariants $\bar{a}^{(m)j,i+2}$ of the weighted homogeneous normal form $\Sigma_{\text{NF}}^{(m)}$, defined by (4.36) and (4.37), are given by

$$\bar{a}^{(m)j,i+2}(x) = \frac{\partial^2 \bar{f}_j^{(m)}}{\partial x_{n-i}^2}(\pi_{n-i}(x))$$

4.7.7 Explicit Normalizing Transformations

In this section, we present the explicit weighted homogeneous transformations bringing the system $\Sigma^{(m)}$ into its normal form $\Sigma_{\text{NF}}^{(m)}$. They have two main advantages: first, they are easily computable (via differentiation and integration of polynomials); secondly, the proof of Theorem 16 giving the normal form $\Sigma_{\text{NF}}^{(m)}$ is based on such transformations.

For any $r+1 \leq i \leq n$ and any $1 \leq j < i \leq n$, the homogeneous polynomial $\psi_{j,i}^{(m-1)}$ is defined by

$$\psi_{j,i}^{(m-1)} = C_j X_{n-i}^{(m-1)}$$

For any $r+1 \leq i \leq j \leq n$, we define recursively the polynomials $\psi_{j,i}^{(m-1)}$ by setting

$$\psi_{j,r}^{(m-1)} = \psi_{r+1,r+1}^{(m-1)} = 0$$

and by taking

$$\psi_{j,i}^{(m-1)} = \frac{\partial f_{j-1}^{(m)}}{\partial \xi_i} + L_{A\xi + f^{(1)}} \psi_{j-1,i}^{(m-1)} + \psi_{j-1,i-1}^{(m-1)} + \int_0^{\xi_i} \frac{\partial \psi_{j-1,i}^{(m-1)}}{\partial \xi_{i-1}} d\xi_i$$

Note that the degree of the homogeneous polynomial $\psi_{j,i}^{(m-1)}$ is either $m-2$ if $1 \leq j \leq r$ or $m-1$ if $r+1 \leq j \leq n$.

Consider the weighted homogeneous feedback transformation

$$\begin{aligned} \Gamma^{(m)} : \quad x &= \xi + \phi^{(m)}(\xi) \\ u &= v + \alpha^{(m)}(\xi) + \beta^{(m-1)}(\xi)v \end{aligned}$$

defined, for any $1 \leq j \leq n$, by

$$\begin{aligned}\phi_j^{(m)}(\xi) &= \sum_{i=r+1}^n \int_0^{\xi_i} \psi_{j,i}^{(m-1)}(\bar{\xi}_i) d\xi_i \\ \alpha^{(m)}(\xi) &= -f_n^{(m)}(\xi) - L_{A\xi+f(1)}\phi_n^{(m)}(\xi) \\ \beta^{(m-1)}(\xi) &= -g_n^{(m-1)}(\xi) - L_B\phi_n^{(m)}(\xi)\end{aligned}\tag{4.38}$$

We have the following result.

THEOREM 18

For any $m \geq 2$, the weighted homogeneous feedback transformation $\Gamma^{(m)}$, defined by (4.38), brings the weighted homogeneous system $\Sigma^{(m)}$, given by (4.33), into its weighted homogeneous normal form $\Sigma_{\text{NF}}^{(m)}$, defined by (4.36).

4.7.8 Weighted Normal Form for Single-Input Systems with Uncontrollable Linearization

In this section, we present our main result giving a normal form under a formal feedback transformation Γ^∞ (see Section 4.3 for some comments on formal feedback) of any single-input control system (with controllable or uncontrollable linearization). For any $1 \leq i \leq n$, we denote $\pi_i(x) = (x_1, \dots, x_i)^T$.

THEOREM 19

Consider the system Σ^∞ , given by (4.27), and assume that all eigenvalues of the uncontrollable part of the linear approximation are real. There exists a formal feedback transformation Γ^∞ of the form (4.28), which brings the system Σ^∞ , given by (4.27), into its normal form

$$\Sigma_{\text{NF}}^\infty : \dot{x} = Ax + Bv + \bar{f}(x)$$

given by

$$\dot{x}_j = \begin{cases} \lambda_j x_j + \sigma_j x_{j+1} + \bar{f}_j(x) & \text{if } 1 \leq j \leq r \\ x_{j+1} + \bar{f}_j(x) & \text{if } r+1 \leq j \leq n-1 \\ v & \text{if } j = n \end{cases}\tag{4.39}$$

where

$$\bar{f}_j(x) = \begin{cases} \sum_{k \in \mathcal{R}_j} \gamma_j^k x_1^{k_1} \cdots x_r^{k_r} + x_{r+1} S_j(\pi(x_{r+1})) \\ \quad + \sum_{i=r+2}^n x_i^2 Q_{j,i}(\pi_i(x)), & \text{if } 1 \leq j \leq r \\ \sum_{i=j+2}^n x_i^2 P_{j,i}(\pi_i(x)), & \text{if } r+1 \leq j \leq n-2 \\ 0, & \text{if } n-1 \leq j \leq n \end{cases} \quad (4.40)$$

where $P_{j,i}$, $Q_{j,i}$, and S_j are formal power series of the indicated variables and $\gamma_j^k \in \mathbb{R}$.

REMARK 1

In the general case of complex eigenvalues of the uncontrollable part, for each complex eigenvalue $\lambda_j = \alpha_j + i\beta_j$, we replace the expression for $\dot{x}_j$ by

$$\begin{pmatrix} \dot{x}_{j,1} \\ \dot{x}_{j,2} \end{pmatrix} = \begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix} \begin{pmatrix} x_{j,1} \\ x_{j,2} \end{pmatrix} + \Xi_j \begin{pmatrix} x_{j,1} \\ x_{j,2} \end{pmatrix} + \bar{f}_j$$

where $\bar{f}_j = (\bar{f}_{j,1}, \bar{f}_{j,2})^T$, for $1 \leq j \leq r$, being defined by formula (4.40), with $\gamma_j^k = (\gamma_{j,1}^k, \gamma_{j,2}^k)^T \in \mathbb{R}^2$, and S_j and $Q_{j,i}$ being $\mathbb{R}^2$ -valued formal power series of the indicated variables. Of course, the resonant set of a complex eigenvalue λ_j is the same as that corresponding to its conjugate $\bar{\lambda}_j$, which explains why we have the same $\mathcal{R}_j$ for both components $x_{j,1}$ and $x_{j,2}$.

Note that the normal form is a natural combination of the two extreme cases: that of dynamical systems and that of systems with controllable linearization. Indeed, if $r = n$, we deal with a dynamical system, then the coordinates $(x_{r+1}, \dots, x_n)$ are not present and the normal form $\Sigma_{\text{NF}}^\infty$ reduces to a dynamical system $\dot{x} = Jx + \bar{f}(x)$ containing resonant terms only, that is, $\bar{f}_j(x) = \sum_{k \in \mathcal{R}_j} \gamma_j^k x_1^{k_1} \cdots x_r^{k_r}$, for $1 \leq j \leq n$. This is, of course, Poincaré normal form of a vector field under a formal diffeomorphism [3] (see also Theorem 2). On the other hand, if $r = 0$ (i.e., the linear approximation is controllable), the coordinates $(x_1, \dots, x_r)$ are not present and our normal form reduces to $\Sigma_{\text{NF}}^\infty$ of Kang [50] (see Section 4.3), for which $\bar{f}_j(x) = \sum_{i=j+2}^n x_i^2 P_{j,i}(\pi_i(x))$, for $1 \leq j \leq n-2$ and $\bar{f}_j(x) = 0$ otherwise.

Another normal form for nonlinear single-input systems with uncontrollable linearization was obtained by the authors earlier [78] (see also [77]) and by Krener et al. [58–60]. Note, however, that those normal forms differ substantially from the one proposed in this paper and in [81]. Indeed, in the approach presented, the homogeneity is calculated with respect to the linearly controllable variables while it is calculated with respect to all variables elsewhere [58–60, 78], compare Example 10.

4.7.9 Example

Example 11 (Kapitsa Pendulum)

In this example, we consider the Kapitsa pendulum whose equations are given by [7, 19]

$$\begin{aligned}\dot{\alpha} &= p + \frac{w}{l} \sin \alpha \\ \dot{p} &= \left(gl - \frac{w^2}{l^2} \cos \alpha \right) \sin \alpha - \frac{w}{l} p \cos \alpha \\ \dot{z} &= w\end{aligned}\tag{4.41}$$

where α is the angle of the pendulum with the vertical z -axis, w is the velocity of the suspension point z , p is proportional to the generalized impulsion, g is the gravity constant, and l is the length of the pendulum.

We assume to control the acceleration $a = \dot{w}$. Introduce the coordinate system $(\xi_1, \xi_2, \xi_3, \xi_4) = (\alpha, p, z/l, w/l)$ and take $u = a/l$ as the control.

The system (4.41) considered around an equilibrium point $(\alpha_0, p_0, z_0, u_0) = (k\pi, 0, 0, 0)$, where $k \in \mathbb{Z}$, rewrites as

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 + \xi_1 \xi_4 + \xi_4 T_1(\xi_1) \\ \dot{\xi}_2 &= \epsilon g_0 \xi_1 - \xi_2 \xi_4 + \xi_2 \xi_4 T_2(\xi_1) + \xi_4^2 Q_2(\xi_1) + R_2(\xi_1) \\ \dot{\xi}_3 &= \xi_4 \\ \dot{\xi}_4 &= u\end{aligned}\tag{4.42}$$

where $g_0 = g/l$; $\epsilon = \pm 1$; T_1, T_2 , and R_2 are analytic functions whose 1-jets at $(k\pi, 0, 0, 0)$ vanish; and Q_2 is an analytic function vanishing at $(k\pi, 0, 0, 0)$. The case $\epsilon = 1$ corresponds to $\alpha_0 = 2n\pi$ and the case $\epsilon = -1$ to $\alpha_0 = (2n + 1)\pi$. One can easily check that the quadratic feedback transformation

$$\begin{aligned}y_1 &= \xi_1 - \xi_1 \xi_3 \\ \Gamma^2 : \quad y_2 &= \xi_2 + \xi_2 \xi_3 \\ y_3 &= \xi_3 \\ y_4 &= \xi_4\end{aligned}$$

brings the system (4.42) into the system

$$\begin{aligned}
 \dot{y}_1 &= y_2 + y_2 y_3 \tilde{S}_1(y_3) + y_4 \tilde{T}_1(y_1, y_3) \\
 \dot{y}_2 &= \varepsilon g_0 y_1 + y_1 y_3 \tilde{S}_2(y_1, y_3) + y_4 \tilde{T}_2(y_1, y_2, y_3) + y_4^2 \tilde{Q}_2(y_1, y_3) + \tilde{R}_2(y_1) \\
 \dot{y}_3 &= y_4 \\
 \dot{y}_4 &= u
 \end{aligned} \tag{4.43}$$

where $\tilde{Q}_1, \tilde{Q}_2, \tilde{R}_2, \tilde{S}_2, \tilde{T}_1$, and $\tilde{T}_2$ are analytic functions.

Since the vector field defined in $\mathbb{R}^3$ by

$$f = \tilde{T}_1(y_1, y_3) \frac{\partial}{\partial y_1} + \tilde{T}_2(y_1, y_2, y_3) \frac{\partial}{\partial y_2} + \frac{\partial}{\partial y_3}$$

does not vanish at $(0, 0, 0) \in \mathbb{R}^3$ (resp. at $(\pi, 0, 0) \in \mathbb{R}^3$), there exists, in a neighborhood of $(0, 0, 0) \in \mathbb{R}^3$ (resp. of $(\pi, 0, 0) \in \mathbb{R}^3$), an analytic transformation $x = \phi(y)$ of the form

$$\begin{aligned}
 x_1 &= \phi_1(y_1, y_3) \\
 x_2 &= \phi_2(y_1, y_2, y_3) \\
 x_3 &= y_3
 \end{aligned}$$

such that

$$(\phi_* f)(x) = \frac{\partial}{\partial x_3}$$

This transformation, completed with $x_4 = y_4$ and $u = v$, brings the system (4.43) into the normal form (compare with Theorem 19)

$$\begin{aligned}
 \dot{x}_1 &= x_2 + \bar{R}_1(x_1, x_2) + x_3 \bar{S}_1(\pi_3(x)) \\
 \dot{x}_2 &= \varepsilon g_0 x_1 + \bar{R}_2(x_1, x_2) + x_3 \bar{S}_2(\pi_3(x)) + x_4^2 \bar{Q}_2(\pi_3(x)) \\
 \dot{x}_3 &= x_4 \\
 \dot{x}_4 &= v
 \end{aligned} \tag{4.44}$$

where $\pi_3(x) = (x_1, x_2, x_3)$.

Clearly, the dimension of the linearly controllable part of (4.42), (i.e., that of (4.44)), equals 2, which means that $r = 2$.

In the case $\varepsilon = 1$, the eigenvalues of the uncontrollable linear part are $\lambda_1 = -\lambda_2 = \sqrt{g_0}$, while in the case $\varepsilon = -1$, they are given by $\lambda_1 = -\lambda_2 = i\sqrt{g_0}$.

In both cases, those eigenvalues are resonant and satisfy, for any $m \geq 2$, the relations

$$\lambda_1 = m\lambda_1 + (m-1)\lambda_2 \quad \text{and} \quad \lambda_2 = (m-1)\lambda_1 + m\lambda_2$$

Applying Poincaré's method [3], we get rid, by a formal diffeomorphism in the space (x_1, x_2) , of all nonresonant terms of the dynamical system

$$\begin{aligned} \dot{x}_1 &= x_2 + \bar{R}_1(x_1, x_2) \\ \dot{x}_2 &= \varepsilon g_0 x_1 + \bar{R}_2(x_1, x_2) \end{aligned}$$

and thus we transform the system, for $\varepsilon = 1$ and $\varepsilon = -1$, into one of the following normal forms.

Set $\lambda = \sqrt{g_0}$. For $\varepsilon = 1$, which is the case of real eigenvalues, the normal form is given by (compare with Theorem 19)

$$\begin{aligned} \dot{x}_1 &= \lambda x_1 + \sum_{m=2}^{\infty} a_m x_1 (x_1 x_2)^{m-1} + x_3 \bar{S}_1(\pi_3(x)) + x_4^2 \bar{Q}_1(\pi_3(x)) \\ \dot{x}_2 &= -\lambda x_2 + \sum_{m=2}^{\infty} b_m x_2 (x_1 x_2)^{m-1} + x_3 \bar{S}_2(\pi_3(x)) + x_4^2 \bar{Q}_2(\pi_3(x)) \end{aligned} \quad (4.45)$$

$$\dot{x}_3 = x_4$$

$$\dot{x}_4 = v$$

For $\varepsilon = -1$, which corresponds to the case of complex eigenvalues (see Remark 1), the normal form is given by

$$\begin{aligned} \dot{x}_1 &= \lambda x_2 + \sum_{m=2}^{\infty} (c_m x_1 + d_m x_2) \left(x_1^2 + x_2^2 \right)^{m-1} + x_3 \tilde{S}_1(\pi_3(x)) + x_4^2 \tilde{Q}_1(\pi_3(x)) \\ \dot{x}_2 &= -\lambda x_1 + \sum_{m=2}^{\infty} (-d_m x_1 + c_m x_2) \left(x_1^2 + x_2^2 \right)^{m-1} + x_3 \tilde{S}_2(\pi_3(x)) + x_4^2 \tilde{Q}_2(\pi_3(x)) \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= v \end{aligned} \quad (4.46)$$

The functions $\bar{S}_1$, $\bar{Q}_1$, $\bar{S}_2$, and $\bar{Q}_2$ on the one hand, and the functions $\tilde{S}_1$, $\tilde{Q}_1$, $\tilde{S}_2$, and $\tilde{Q}_2$ on the other, are formal power series which, in general, are different from the objects denoted earlier by the same symbols.

Note that we transform the original system into its normal form (4.44) using analytic feedback transformations and that the only passage defined by a formal feedback transformation is that transforming the system (4.44) into (4.45) (resp. (4.46)).

4.8 Normal Forms for Multi-Input Nonlinear Control Systems

In this section, we present a generalization of normal forms obtained in Section 4.3 for multi-input nonlinear control systems with controllable linearization [88]. Normal forms for two-input nonlinear control systems have been obtained previously [86], and will be derived here as a particular case. The general case of multi-input systems with uncontrollable linearization will appear elsewhere [89]. We will illustrate normal forms in this section by considering three physical examples: a model of a crane in Example 12, a model of a planar vertical takeoff and landing aircraft in Example 13, and finally prototype of a wireless multi-vehicle testbed in Example 14 [15, 17].

Consider control systems of the form

$$\Pi : \dot{\xi} = F(\xi, u), \quad \xi \in \mathbb{R}^n, u = (u_1, \dots, u_p)^T \in \mathbb{R}^p$$

around the equilibrium point $(0, 0) \in \mathbb{R}^n \times \mathbb{R}^p$, that is, $f(0, 0) = 0$, and denote by

$$\Pi^{[1]} : \dot{\xi} = \mathbb{F}\xi + \mathbb{G}u = \mathbb{F}\xi + \mathbb{G}_1 u_1 + \dots + \mathbb{G}_p u_p$$

its linearization at this point, where

$$\mathbb{F} = \frac{\partial F}{\partial \xi}(0, 0), \quad \mathbb{G}_1 = \frac{\partial F}{\partial u_1}(0, 0), \dots, \quad \mathbb{G}_p = \frac{\partial F}{\partial u_p}(0, 0)$$

We will assume for simplicity [88, 89] that $\mathbb{G}_1 \wedge \dots \wedge \mathbb{G}_p \neq 0$, and the linearization is controllable, that is

$$\text{span}\{\mathbb{F}^i \mathbb{G}_k : 0 \leq i \leq n-1, 1 \leq k \leq p\} = \mathbb{R}^n$$

Let $(r_1, \dots, r_p)$, $1 \leq r_1 \leq \dots \leq r_p = r$, be the largest, in the lexicographic ordering, p -tuple of positive integers, with $r_1 + \dots + r_p = n$, such that

$$\text{span}\{\mathbb{F}^i \mathbb{G}_k : 0 \leq i \leq r_k - 1, 1 \leq k \leq p\} = \mathbb{R}^n \quad (4.47)$$

Without loss of generality we can assume that the linearization is in Brunovský canonical form

$$\Pi_{\text{CF}}^{[1]} : \dot{\xi} = A\xi + Bu = A\xi + B_1u_1 + \cdots + B_pu_p$$

where $A = \text{diag}(A_1, \dots, A_p)$, $B = (B_1, \dots, B_p) = \text{diag}(b_1, \dots, b_p)$, that is,

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_p \end{pmatrix}_{n \times n}, \quad B = \begin{pmatrix} b_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_p \end{pmatrix}_{n \times p} \quad (4.48)$$

with (A_k, b_k) in Brunovský single-input canonical forms of dimensions r_k , for any $1 \leq k \leq p$.

With the p -tuple $(r_1, \dots, r_p)$, we associate the p -tuple $(d_1, \dots, d_p)$ of nonnegative integers, $0 = d_p \leq \cdots \leq d_1 \leq r - 1$, such that $r_1 + d_1 = \cdots = r_p + d_p = r$.

Our aim is to give a normal form of feedback classification of such systems under invertible feedback transformations of the form

$$\Upsilon : \begin{aligned} x &= \phi(\xi) \\ u &= \psi(\xi, v) \end{aligned}$$

where $\phi(0) = 0$ and $\psi(0, 0) = 0$.

We study, step by step, the action of the Taylor series expansion Υ^∞ of the feedback transformation Υ , given by

$$\begin{aligned} \Upsilon^\infty : \quad x &= \phi(\xi) = \xi + \sum_{m=2}^{\infty} \phi^{[m]}(\xi) \\ u &= \psi(\xi, v) = v + \sum_{m=2}^{\infty} \psi^{[m]}(\xi, v) \end{aligned} \quad (4.49)$$

on the Taylor series expansion Π^∞ of the system Π , given by

$$\Pi^\infty : \dot{\xi} = A\xi + Bu + \sum_{m=2}^{\infty} f^{[m]}(\xi, u) \quad (4.50)$$

Throughout this section, in particular, in formulas (4.49) and (4.50), the homogeneity of $f^{[m]}$ and $\psi^{[m]}$ will be taken with respect to the variables ξ, v and ξ, u , respectively.

4.8.1 Non-affine Normal Forms

Let $1 \leq s \leq t \leq p$. We denote by

$$x_s = (x_{s,d_s+1}, \dots, x_{s,r}), \quad x_{s,r+1} = v_s$$

and we set $\bar{x}_{s,i} = (x_{s,d_s+1}, \dots, x_{s,i})$ for any $d_s + 1 \leq i \leq r + 1$.

We also define the projections

$$\pi_{t,i}^s(x) = (\bar{x}_{1,i}, \dots, \bar{x}_{s,i}, \bar{x}_{s+1,i-1}, \dots, \bar{x}_{t-1,i-1}, \bar{x}_{t,i}, \bar{x}_{t+1,i-1}, \dots, \bar{x}_{p,i-1})$$

where $\bar{x}_{s,i}$ is empty whenever $0 \leq i \leq d_s$.

Our main result for multi-input nonlinear control systems with controllable linearization is as follows.

THEOREM 20

The control system Π^∞ , defined by (4.50), is feedback equivalent, by a formal feedback transformation Υ^∞ of the form (4.49), to the normal form

$$\Pi_{\text{NF}}^\infty : \dot{x} = Ax + Bv + \sum_{m=2}^{\infty} \bar{f}^{[m]}(x, v)$$

where for any $m \geq 2$, we have

$$\bar{f}^{[m]}(x, v) = \sum_{k=1}^p \sum_{j=d_k+1}^{r-1} \bar{f}_j^{k[m]}(x, v) \frac{\partial}{\partial x_{k,j}} \quad (4.51)$$

with

$$\begin{aligned} \bar{f}_j^{k[m]}(x, v) = & \sum_{1 \leq s \leq t \leq p} \sum_{i=j+2}^{r+1} x_{s,i} x_{t,i} P_{j,i,s,t}^{k[m-2]} (\pi_{t,i}^s(x)) \\ & + \sum_{1 \leq s < t \leq p} \sum_{i=j+2}^{r+1} x_{s,i} x_{t,i-1} Q_{j,i,s,t}^{k[m-2]} (\pi_{t,i-1}^s(x)) \end{aligned} \quad (4.52)$$

for any $1 \leq k \leq p$ and any $d_k + 1 \leq j \leq r - 1$.

The functions $P_{j,i,s,t}^{k[m-2]}$ and $Q_{j,i,s,t}^{k[m-2]}$ stand for homogeneous polynomials of degree $m - 2$ of the indicated variables; $P_{j,i,s,t}^{k[m-2]}$ (resp. $Q_{j,i,s,t}^{k[m-2]}$) being equal to zero for $1 \leq i \leq d_s$ (resp. $1 \leq i \leq d_s + 1$).

Notice that when $p = 1$, that is, if we deal with single-input control systems, then the homogeneous polynomials $Q_{j,i,s,t}^{k[m-2]}$ are zero, $v = n$, and thus the normal form reduces to the non-affine of Kang normal form given by

$$\bar{f}_j^{[m]}(x, v) = \sum_{i=j+2}^{r+1} x_{1,i}^2 P_{j,i,1,1}^{k[m-2]}(\pi_{1,i}^1(x)) = \sum_{i=j+2}^{r+1} x_i^2 P_{j,i}^{[m-2]}(\bar{x}_i)$$

Two-input control systems: the normal form for two-input control systems with controllable linearization was obtained before [86] and deduces as the following corollary from Theorem 20 by taking $p = 2$.

COROLLARY 3

The control system Π^∞ , defined by (4.50) with $p = 2$, is feedback equivalent, by a formal feedback transformation Υ^∞ of the form (4.49), to the normal form

$$\Pi_{\text{NF}}^\infty : \dot{x} = Ax + Bv + \sum_{m=2}^{\infty} \bar{f}^{[m]}(x, v)$$

where for any $m \geq 2$, we have

$$\bar{f}_j^{[m]}(x, v) = \sum_{j=d_1+1}^{r-1} \bar{f}_j^{1[m]}(x, v) \frac{\partial}{\partial x_{1,j}} + \sum_{j=d_2+1}^{r-1} \bar{f}_j^{2[m]}(x, v) \frac{\partial}{\partial x_{2,j}}$$

with,

$$\begin{aligned} \bar{f}_j^{k[m]}(x, v) = & \sum_{i=j+2}^{r+1} \left(x_{1,i}^2 P_{j,i}^{k[m-2]}(\bar{x}_{1,i}, \bar{x}_{2,i-1}) + x_{2,i}^2 Q_{j,i}^{k[m-2]}(\bar{x}_{1,i-1}, \bar{x}_{2,i}) \right) \\ & + \sum_{i=j+2}^r \left(x_{1,i} x_{2,i} R_{j,i}^{k[m-2]}(\bar{x}_{1,i}, \bar{x}_{2,i}) + x_{1,i} x_{2,i-1} S_{j,i}^{k[m-2]}(\bar{x}_{1,i-1}, \bar{x}_{2,i-1}) \right) \end{aligned}$$

for any $k = 1, 2$ and any $d_k + 1 \leq j \leq r - 1$.

The homogeneous polynomials $P_{j,i}^{k[m-2]}$, $Q_{j,i}^{k[m-2]}$, and $S_{j,i}^{k[m-2]}$ being equal to zero for $1 \leq i \leq d_1$.

4.8.2 Affine Normal Forms

Here we consider the action of Γ^∞ , given by

$$\begin{aligned} x &= \phi(\xi) = T\xi + \sum_{m=2}^{\infty} \phi^{[m]}(\xi) \\ \Gamma^\infty : \\ u &= \alpha(\xi) + \beta(\xi)v = K\xi + Lv + \sum_{m=2}^{\infty} \left(\alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v \right) \end{aligned} \quad (4.53)$$

on multi-input Σ^∞ , given by

$$\Sigma^\infty : \dot{\xi} = A\xi + Bu + \sum_{m=2}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right) \quad (4.54)$$

where we assume the linear part to be already in the Brunovský canonical form (A, B) (see (4.48)).

The main result of [88] in the affine case is next described.

THEOREM 21

1. The formal system Σ^∞ , defined by (4.54), is feedback equivalent, by a formal feedback transformation Γ^∞ of the form (4.53), to the normal form

$$\begin{aligned} \Sigma_{\text{NF}}^\infty : \dot{x} &= Ax + Bv \\ &+ \sum_{m=2}^{\infty} \left(\bar{f}^{[m]}(x) + \bar{g}_1^{[m-1]}(x)v_1 + \cdots + \bar{g}_{p-1}^{[m-1]}(x)v_{p-1} \right) \end{aligned}$$

where for any $m \geq 2$, the vector field $\bar{f}^{[m]}(x)$ and the vector fields $\bar{g}_s^{[m-1]}(x)$, for $1 \leq s \leq p-1$, are given by

$$\begin{aligned} \bar{f}^{[m]}(x) &= \sum_{k=1}^p \sum_{j=d_k+1}^{r-1} \bar{f}_j^{k[m]}(x) \frac{\partial}{\partial x_{k,j}}, \\ \bar{g}_s^{[m-1]}(x) &= \sum_{k=1}^p \sum_{j=d_k+1}^{r-1} \bar{g}_{s,j}^{k[m-1]}(x) \frac{\partial}{\partial x_{k,j}} \end{aligned}$$

with

$$\begin{aligned}\tilde{f}_j^{k[m]}(x) &= \sum_{1 \leq s \leq t \leq p} \sum_{i=j+2}^r x_{s,i} x_{t,i} P_{j,i,s,t}^{k[m-2]} (\pi_{t,i}^s(x)) \\ &\quad + \sum_{1 \leq s < t \leq p} \sum_{i=j+2}^r x_{s,i} x_{t,i-1} Q_{j,i,s,t}^{k[m-2]} (\pi_{t,i-1}^s(x))\end{aligned}$$

and

$$\tilde{g}_{s,j}^{k[m-1]}(x) = \sum_{t=s+1}^p x_{t,r} Q_{j,r+1,s,t}^{k[m-2]} (\pi_{t,r}^s(x))$$

for any $1 \leq k \leq p$ and any $d_k + 1 \leq j \leq r - 1$.

2. Moreover, if the formal distribution

$$\mathcal{G}^\infty = \text{span} \left\{ B_1 + \sum_{m=2}^{\infty} g_1^{[m-1]}, \dots, B_p + \sum_{m=2}^{\infty} g_p^{[m-1]} \right\}$$

is involutive, then the homogeneous polynomials $Q_{j,r+1,s,t}^{k[m-2]}$ are equal to zero, that is, the normal form reduces to

$$\Sigma_{\text{NF}}^\infty : \dot{x} = Ax + Bv + \sum_{m=2}^{\infty} \tilde{f}^{[m]}(x)$$

REMARK 2

Note that only $p - 1$ control vector fields are present in the normal form, the control vector field g_p being normalized to $(0, \dots, 0, 1)^T$. This is what happens in the single-input case. If we take $p = 1$, then all homogeneous control vector fields $\tilde{g}_s^{[m-1]}(x)$, as well as the homogeneous polynomials $Q_{j,i,s,t}^{k[m-2]}(\pi_{t,i-1}^s(x))$ are not present in the aforementioned normal form. Thus, the normal form given by Theorem 21 reduces to the Kang normal form. In item (2) we rediscover a well-known result: if a nonsingular distribution is involutive, there are coordinates that normalize the whole distribution to a constant one.

4.8.3 Examples

In this section, we will illustrate the theory of normal forms for multi-input control systems by considering three physical examples. We will first treat the case of a model of a crane, then a prototype of a planar vertical

takeoff and landing aircraft, and finally we will consider the model of a multi-vehicle wireless testbed presented by Caltech.

Example 12 (Model of a crane)

Consider the following model of a crane borrowed from earlier work [16, 18]. The state equations are

$$\begin{aligned}\dot{z}_1 &= z_2 \\ \dot{z}_2 &= -\frac{g \sin z_1}{R} - \frac{2z_2}{R}\dot{R} - d\frac{\cos z_1}{R}\end{aligned}$$

where z_1 is the angle between the rope and the vertical axis, z_2 the angular velocity, R the length of the rope, and d the trolley acceleration. We aim to control $\dot{R} = u_1$ and $\dot{d} = u_2$ (notice that in Ref. [18] controls are R and D , where $d = \ddot{D}$). We consider the system around the equilibrium point $z_{10} = z_{20} = d_0 = 0$, $R_0 = 1$. The linear approximation is controllable with controllability indices $r_1 = 1$ and $r_2 = 3$. Then $d_1 = 2$ and $d_2 = 0$. Introduce the coordinates

$$\begin{aligned}\xi_{1,3} &= R - R_0 = R - 1 \\ \xi_{2,1} &= z_1 \\ \xi_{2,2} &= z_2 \\ \xi_{2,3} &= d\end{aligned}$$

in which the system takes the form

$$\begin{aligned}\dot{\xi}_{1,3} &= u_1 \\ \dot{\xi}_{2,1} &= \xi_{2,2} \\ \dot{\xi}_{2,2} &= -\frac{g \sin \xi_{2,1}}{1 + \xi_{1,3}} - \xi_{2,3} \frac{\cos \xi_{2,1}}{1 + \xi_{1,3}} - \frac{2\xi_{2,2}}{1 + \xi_{1,3}} u_1 \\ \dot{\xi}_{2,3} &= u_2\end{aligned}$$

To bring this system to its normal form, we rectify the involutive distribution $\mathcal{G} = \text{span}\{g_1, g_2\}$, with $g_1 = (1, 0, -\frac{2\xi_{2,2}}{1+\xi_{1,3}}, 0)^T$ and $g_2 = (0, 0, 0, 1)^T$, and we normalize the component $f_{2,2}^2$ by taking

$$\begin{aligned}x_{1,3} &= \xi_{1,3} \\ x_{2,1} &= \xi_{2,1} \\ x_{2,2} &= \xi_{2,2}(1 + \xi_{1,3})^2 \\ x_{2,3} &= -(1 + \xi_{1,3})(g \sin \xi_{2,1} + \xi_{2,3} \cos \xi_{2,1})\end{aligned}$$

followed by a suitable feedback. This yields

$$\begin{aligned}\dot{x}_{1,3} &= u_1 \\ \dot{x}_{2,1} &= x_{2,2} - 2x_{1,3}x_{2,2} + 3x_{1,3}^2 \frac{x_{2,2}}{(1+x_{1,3})^2} \\ \dot{x}_{2,2} &= x_{2,3} \\ \dot{x}_{2,3} &= u_2\end{aligned}$$

which is in a normal form (compare with Corollary 3), where all homogeneous polynomials $P_{j,i}^{k[m-2]}$, $Q_{j,i}^{k[m-2]}$, $R_{j,i}^{k[m-2]}$, and $S_{j,i}^{k[m-2]}$ are equal to zero except for $S_{1,1}^{2[0]} = -2$ and the homogeneous polynomials $P_{1,1}^{2[m-2]}$, for $m \geq 3$, which are equal to the homogeneous parts of degree $m-2$ of the function $3x_{2,2}/(1+x_{1,3})^2$.

Example 13 (PVTOL aircraft)

In this example, we consider the prototype of a planar vertical takeoff and landing (PVTOL) aircraft. The equations of motion of the PVTOL [92] are given by

$$\begin{aligned}\ddot{x} &= -\sin \theta u_1 + \varepsilon^2 \cos \theta u_2 \\ \ddot{y} &= \cos \theta u_1 + \varepsilon^2 \sin \theta u_2 - 1 \\ \ddot{\theta} &= u_2\end{aligned}$$

where (x, y) denotes the position of the center mass of the aircraft, θ the angle of the aircraft relative to the x -axis, “ -1 ” the gravitational acceleration, and $\varepsilon \neq 0$ the (small) coefficient giving the coupling between the rolling moment and the lateral acceleration of the aircraft. The control inputs u_1 and u_2 are the thrust (directed out the bottom of the aircraft) and the rolling moment.

We introduce the variables

$$\begin{aligned}\xi_{1,1} &= y, & \xi_{1,2} &= \dot{y} \\ \xi_{2,1} &= x, & \xi_{2,2} &= \dot{x} \\ \xi_{2,3} &= \theta, & \xi_{2,4} &= \dot{\theta} \\ w_1 &= u_1 - 1, & w_2 &= u_2.\end{aligned}$$

The equations of motion of the PVTOL become

$$\begin{aligned}
 \dot{\xi}_{1,1} &= \xi_{1,2} \\
 \dot{\xi}_{1,2} &= \cos \xi_{2,3} w_1 + \varepsilon^2 \sin \xi_{2,3} w_2 + \cos \xi_{2,3} - 1 \\
 \dot{\xi}_{2,1} &= \xi_{2,2} \\
 \dot{\xi}_{2,2} &= -\sin \xi_{2,3} w_1 + \varepsilon^2 \cos \xi_{2,3} w_2 - \sin \xi_{2,3} \\
 \dot{\xi}_{2,3} &= \xi_{2,4} \\
 \dot{\xi}_{2,4} &= w_2
 \end{aligned} \tag{4.55}$$

The equilibria of the system are defined by

$$(\xi_{1,1}^e, \xi_{1,2}^e, \xi_{2,1}^e, \xi_{2,2}^e, \xi_{2,3}^e, \xi_{2,4}^e, w_1^e, w_2^e)^T = (c, 0, 0, 0, 0, 0, 0, 0)^T$$

where c is any constant. The linearization of the system (4.55) around the equilibria is given by

$$\begin{aligned}
 \dot{\xi}_{1,1} &= \xi_{1,2} \\
 \dot{\xi}_{1,2} &= w_1 \\
 \dot{\xi}_{2,1} &= \xi_{2,2} \\
 \dot{\xi}_{2,2} &= -\xi_{2,3} + \varepsilon^2 w_2 \\
 \dot{\xi}_{2,3} &= \xi_{2,4} \\
 \dot{\xi}_{2,4} &= w_2
 \end{aligned}$$

It is easy to see that the linear system is controllable with controllability indices $r_1 = 2$ and $r_2 = 4$, and hence $d_1 = 2$ and $d_2 = 0$.

The change of coordinates given by

$$\begin{aligned}
 x_{1,3} &= \xi_{1,1} - \varepsilon^2 \int_0^{\xi_{2,3}} \frac{dt}{\cos t} \\
 x_{1,4} &= \xi_{1,2} + \xi_{1,2} \tan \xi_{2,3} - \frac{\varepsilon^2}{\cos \xi_{2,3}} \xi_{2,4} \\
 x_{2,1} &= \xi_{2,1} \\
 x_{2,2} &= \xi_{2,2} \\
 x_{2,3} &= -\tan \xi_{2,3} \\
 x_{2,4} &= -\xi_{2,4}(1 + \tan^2 \xi_{2,3}) = \dot{x}_{2,3}
 \end{aligned}$$

followed by the feedback

$$w_1 = \frac{v_1}{\cos \xi_{2,1}} - \varepsilon^2 v_2 \tan \xi_{2,1} + \frac{1}{\cos \xi_{2,1}} - 1$$

$$v_2 = \dot{x}_{2,4} = -w_2(1 + \tan^2 \xi_{2,3}) - 2\xi_{2,4}^2 \tan \xi_{2,3}(1 + \tan^2 \xi_{2,3})$$

takes the system into the following one

$$\begin{aligned}\dot{x}_{1,3} &= x_{1,4} \\ \dot{x}_{1,4} &= v_1 \\ \dot{x}_{2,1} &= x_{2,2} + x_{1,4}x_{2,3} \\ \dot{x}_{2,2} &= x_{2,3} - x_{1,4}x_{2,4} + \varepsilon^2 (1 - x_{2,3}^2) x_{2,4}^2 \\ \dot{x}_{2,3} &= x_{2,4} \\ \dot{x}_{2,4} &= \tilde{v}_2\end{aligned}$$

This system is in normal form (compare with Corollary 3), with

$$Q_{1,4,1,2}^{2[0]}(x) = 1, \quad P_{2,4,1,2}^{2[0]}(x) = -1, \quad P_{2,4,2,2}^{2[0]}(x) = \varepsilon^2, \quad P_{2,4,2,2}^{2[2]}(x) = -\varepsilon^2 x_{2,3}^2$$

Example 14 (Multi-Vehicle Wireless Testbed)

We consider the Caltech Multi-Vehicle Wireless Testbed, presented elsewhere [15, 17] and we study the normal form of one vehicle. The equations of motion of an MVWT vehicle [15, 17] are given by

$$\begin{aligned}m\ddot{x} &= -\eta\dot{x} + (F_s + F_p) \cos \theta \\ m\ddot{y} &= -\eta\dot{y} + (F_s + F_p) \sin \theta \\ J\ddot{\theta} &= -\psi\dot{\theta} + (F_s - F_p)l\end{aligned}$$

where (x, y) denotes the position of the center mass of the vehicle, θ the angle of the axis of the vehicle with the x -axis, m the mass of the vehicle, J the rotational inertia, F_s and F_p denote, respectively, the starboard and port fan forces, and l (r in [15, 17]) the common moment arm of the forces. The center mass of the vehicle and the center of geometry are assumed to coincide. The constants η and ψ stand, respectively, for the coefficients of viscous friction and rotational friction.

Let us introduce the variables

$$\begin{aligned}\xi_{0,1} &= y, & \xi_{0,2} &= \dot{y}, & \xi_{1,1} &= x, & \xi_{1,2} &= \dot{x}, \\ \xi_{2,1} &= \theta, & \xi_{2,2} &= \dot{\theta}, & u_1 &= F_s + F_p, & u_2 &= F_s - F_p.\end{aligned}$$

The equations of motion of an MVWT vehicle can be rewritten as

$$\begin{aligned}
 \dot{\xi}_{0,1} &= \xi_{0,2} \\
 \dot{\xi}_{0,2} &= -\eta \xi_{0,2} + u_1 \sin \xi_{2,1} \\
 \dot{\xi}_{1,1} &= \xi_{1,2} \\
 \dot{\xi}_{1,2} &= -\eta \xi_{1,2} + u_1 \cos \xi_{2,1} \\
 \dot{\xi}_{2,1} &= \xi_{2,2} \\
 \dot{\xi}_{2,2} &= -\psi \xi_{2,2} + u_2 l
 \end{aligned} \tag{4.56}$$

We can notice that the system is affine and its distribution $\mathcal{G} = \text{span}\{g_1, g_2\}$, where

$$g_1 = (0, \sin \xi_{2,1}, 0, \cos \xi_{2,1}, 0, 0)^T \quad \text{and} \quad g_2 = (0, 0, 0, 0, 0, 1)^T$$

is involutive and of constant rank 2. An equilibrium point for the system (4.56) is any constant position and orientation $(x_c, y_c, \theta_c)^T = (\xi_{1,1}^e, \xi_{0,1}^e, \xi_{2,1}^e)^T$. The linearization of the system (4.56) around an equilibrium (we assume $\theta_c = 0$) is given by

$$\begin{aligned}
 \dot{\xi}_{0,1} &= \xi_{0,2} \\
 \dot{\xi}_{0,2} &= -\eta \xi_{0,2} \\
 \dot{\xi}_{1,1} &= \xi_{1,2} \\
 \dot{\xi}_{1,2} &= -\eta \xi_{1,2} + u_1 \\
 \dot{\xi}_{2,1} &= \xi_{2,2} \\
 \dot{\xi}_{2,2} &= -\psi \xi_{2,2} + u_2 l
 \end{aligned}$$

It is easy to see that this linear system is not controllable because

$$\text{span}\{F^i G_k, 0 \leq i \leq 5, 1 \leq k \leq 2\} = \mathbb{R}^4$$

where

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -\eta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\eta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -\psi \end{pmatrix}, \quad G_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad G_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

The feedback transformation defined by

$$u_1 = \frac{1}{\cos \xi_{2,1}} v_1 + \eta \frac{\xi_{1,2}}{\cos \xi_{2,1}} \quad \text{and} \quad v_2 = \frac{u_2}{l} + \frac{\psi}{l} \xi_{2,2}$$

takes the system into the following one

$$\begin{aligned} \dot{\xi}_{0,1} &= \xi_{0,2} \\ \dot{\xi}_{0,2} &= -\eta \xi_{0,2} + \eta \xi_{1,2} \tan \xi_{2,1} + u_1 \tan \xi_{2,1} \\ \dot{\xi}_{1,1} &= \xi_{1,2} \\ \dot{\xi}_{1,2} &= v_1 \\ \dot{\xi}_{2,1} &= \xi_{2,2} \\ \dot{\xi}_{2,2} &= v_2 \end{aligned}$$

The change of coordinates given by

$$\begin{aligned} x_{0,1} &= \xi_{0,1} \\ x_{0,2} &= \xi_{0,2} - \xi_{1,2} \tan \xi_{2,1} \\ x_{1,1} &= \xi_{1,1} \\ x_{1,2} &= \xi_{1,2} \\ x_{2,1} &= \xi_{2,1} \\ x_{2,2} &= \xi_{2,2} \end{aligned}$$

brings the system into the normal form

$$\begin{aligned} \dot{x}_{0,1} &= x_{0,2} + x_{1,2} \tan x_{2,1} \\ \dot{x}_{0,2} &= -\eta x_{0,2} - x_{1,2} x_{2,2} (1 + \tan^2 x_{2,1}) \\ \dot{x}_{1,1} &= x_{1,2} \\ \dot{x}_{1,2} &= v_1 \\ \dot{x}_{2,1} &= x_{2,2} \\ \dot{x}_{2,2} &= v_2 \end{aligned}$$

Here, the linearly controllable part is feedback linearizable because the indices of controllability are $r_1 = r_2 = 2$.

4.9 Feedback Linearization

In this section, we will discuss the feedback linearization problem. We will recall the well-known result characterizing feedback linearizability in terms of involutivity of certain distributions and compare it with a condition using the homogeneous m -invariants. Then we will consider systems with uncontrollable linearization and, similarly, we will compare a geometric condition involving the involutivity of suitable distributions with a condition using weighted homogeneous invariants. Finally, we will discuss feedback linearizability of general systems (i.e., not necessarily affine in controls).

Consider a C^∞ -smooth control-affine system of the form

$$\Sigma : \dot{\xi} = f(\xi) + \sum_{i=1}^m g_i(\xi)u_i$$

where $f(\xi_0) = 0$, which we will assume throughout this section. To state a feedback linearization result for Σ , we define the following distributions

$$\begin{aligned}\mathcal{D}^1(x) &= \text{span}\{g_i(x), 1 \leq i \leq m\} \\ \mathcal{D}^j(x) &= \text{span}\{\text{ad}_f^{q-1}g_i(x), 1 \leq q \leq j, 1 \leq i \leq m\},\end{aligned}$$

for $j \geq 2$. If the dimensions $m_j(x)$ of $\mathcal{D}^j(x)$ are constant (see (FL1) and (FL1)') we denote them by m_j and we define indices ρ_j as follows. Define $m_0 = 0$ and put $q_j = m_j - m_{j-1}$ for $1 \leq j \leq n$. Then we define

$$\rho_i = \max\{q_j \mid q_j \geq i\} \quad (4.57)$$

Clearly, we have $\rho_1 \geq \dots \geq \rho_m$ (and also $\sum_{i=1}^m \rho_i = n$ if the linear part (F, G) of Σ is controllable). For the linear controllable system $\dot{\xi} = F\xi + Gu$, the integers ρ_i 's form the set of controllability indices (compare Example 1).

We will be interested in feedback linearization, that is, in feedback equivalence of Σ to a linear system of the form

$$\dot{x} = Ax + Bv = Ax + \sum_{i=1}^m b_i v_i$$

under a feedback transformation $x = \phi(\xi)$, $u = \alpha(\xi) + \beta(\xi)v$. The following result [36, 37, 45], see also [38, 66] describes control-affine systems that are locally feedback equivalent to linear controllable systems.

THEOREM 22

The following conditions are equivalent:

1. Σ is locally, at $\xi_0 \in \mathbb{R}^n$, feedback equivalent to a linear controllable system
2. Σ satisfies in a neighborhood of ξ_0 :
 - (FL1) $\dim \mathcal{D}^j(\xi) = \text{const}$, for $1 \leq j \leq n$
 - (FL2) the distributions $\mathcal{D}^j$ are involutive, for $1 \leq j \leq n$
 - (FL3) $\dim \mathcal{D}^n(\xi_0) = n$
3. Σ satisfies in a neighborhood of ξ_0 :
 - (FL1)' $\dim \mathcal{D}^j(\xi) = \text{const}$, for $1 \leq j \leq n$
 - (FL2)' the distributions $\mathcal{D}^{\rho_j-1}$ are involutive, for $1 \leq j \leq m$
 - (FL3)' $\dim \mathcal{D}^{\rho_1}(\xi_0) = n$, where ρ_1 is the largest controllability index

The conditions (FL1)'–(FL3)' involve the minimal number of distributions whose involutivity has to be checked. On the other hand, the conditions (FL1)–(FL3) are more transparent and do not require the calculation of controllability indices.

In the single-input case $m = 1$, the condition (FL3) (or, equivalently, (FL3)') states that $g(\xi_0), \dots, \text{ad}_f^{n-1}g(\xi_0)$ are independent, which implies that all distributions $\mathcal{D}^j$, for $1 \leq j \leq n$, are of constant rank. In the single-input case, we have the following corollary of Theorem 22.

COROLLARY 4

A single-input system Σ is feedback linearizable if and only if it satisfies:

- (FL1)_{SI} $g(\xi_0), \dots, \text{ad}_f^{n-1}g(\xi_0)$ are independent
- (FL2)_{SI} the distribution $\mathcal{D}^{n-1}$ is involutive

Now consider the single input-system Σ , given by $\dot{\xi} = f(\xi) + g(\xi)u$. Without loss of generality we assume that $\xi_0 = 0$. Consider the infinite Taylor series expansion of Σ given at $\xi_0 = 0 \in \mathbb{R}^n$ by

$$\Sigma^\infty : \dot{\xi} = F\xi + Gu + \sum_{m=2}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right)$$

Of course, a necessary condition for feedback equivalence of Σ^∞ to a linear controllable system is that the linear part (F, G) of Σ^∞ is controllable. So we can put it by a linear feedback transformation Γ^1 into the Brunovsky canonical form (A, B) . Now consider the homogeneous system

$$\Sigma^{[m]} : \dot{\xi} = A\xi + Bu + f^{[m]}(\xi) + g^{[m-1]}(\xi)u$$

Recall that $\Delta = \{(j, i) \in \mathbb{N} \times \mathbb{N} : 1 \leq j \leq n-2, 0 \leq i \leq n-j-2\}$ and that, by definition, equivalence of m -homogeneous systems means equivalence modulo terms of degree higher than m (see Section 4.3). Theorem 3 and Theorem 4 imply the following result of Kang [50]:

PROPOSITION 6

The homogenous system $\Sigma^{[m]}$ is equivalent, via a homogeneous feedback transformation Γ^m , to the linear system

$$\dot{x} = Ax + Bv$$

if and only if

$$a^{[m]j,i+2} = 0$$

$(j, i) \in \Delta$, that is, all homogenous m -invariants vanish.

This observation leads to the following linearization step-by-step procedure. Consider the system Σ^∞ and apply a linear feedback transformation Γ^1 bringing the linear part (F, G) into the Brunovský canonical form (A, B) . Denote $\Sigma^{\infty,1} = \Gamma_*^1(\Sigma^\infty)$. Now bring the homogenous system $\Sigma^{[2]}$ of $\Sigma^{\infty,1}$ into its normal form $\Sigma_{\text{NF}}^{[2]}$ via a homogeneous transformation Γ^2 . If the 2-invariants $a^{[2]j,i+2}$, $(j, i) \in \Delta$, vanish then the system

$$\Sigma^{\infty,2} = \Gamma_*^2(\Sigma^{\infty,1})$$

is linear modulo terms in $V^{\geq 3}$. Notice that, although the transformation Γ^2 is determined by the homogenous part $\Sigma^{[2]}$ of $\Sigma^{\infty,1}$ only, we apply Γ^2 to the whole system $\Sigma^{\infty,1}$ and thus we modify, in general, all terms of $\Sigma^{\infty,1}$ to get $\Sigma^{\infty,2}$. Now suppose that a sequence of systems $\Sigma^{\infty,1}, \dots, \Sigma^{\infty,m-1}$ has been defined, and $\Sigma^{\infty,m-1}$ is linear modulo terms in $V^{\geq m}$. Bring the homogenous system $\Sigma^{[m]}$ of $\Sigma^{\infty,m-1}$ into its normal form $\Sigma_{\text{NF}}^{[m]}$ via a homogeneous transformation Γ^m . If the m -invariants $a^{[m]j,i+2}$, $(j, i) \in \Delta$, vanish then the system

$$\Sigma^{\infty,m} = \Gamma_*^m(\Sigma^{\infty,m-1})$$

is linear modulo terms in $V^{\geq m+1}$.

We thus have the following counterpart of Corollary 4.

PROPOSITION 7

The system Σ^∞ , with controllable linearization, is feedback equivalent via a formal feedback Γ^∞ to a controllable linear system $\dot{x} = Ax + Bv$ if and only if

for any $m \geq 2$

$$a^{[m]j,i+2} = 0$$

where $(j, i) \in \Delta$ and $a^{[m]j,i+2}$ are m -invariants of the homogeneous system $\Sigma^{[m]}$ of $\Sigma^{\infty, m-1} = \Gamma_*^{m-1} \dots \Gamma_*^2 \Gamma_*^1(\Sigma^\infty)$.

Clearly, if a system Σ is feedback linearizable (i.e., satisfies the conditions (FL1)_{SI} and (FL2)_{SI} of Corollary 4), then its infinite Taylor series expansion Σ^∞ satisfies the conditions of Proposition 7. Now we will answer the important question of whether we can reverse this implication.

PROPOSITION 8

Consider an analytic system Σ , with a controllable linearization (F, G) . Assume that

$$a^{[m]j,i+2} = 0, \quad \text{for any } m \geq 2$$

where $(j, i) \in \Delta$ and $a^{[m]j,i+2}$ are m -invariants of the homogeneous system $\Sigma^{[m]}$ of $\Sigma^{\infty, m-1} = \Gamma_^{m-1} \dots \Gamma_*^2 \Gamma_*^1(\Sigma^\infty)$. Then Σ is equivalent, via a local analytic feedback transformation Γ , to a controllable linear system $\dot{x} = Ax + Bv$.*

An analogous result does not hold in the C^∞ -category. To see this result consider, for example, the system

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 + f_1(\xi_n) \\ \dot{\xi}_2 &= \xi_3 \\ &\vdots \\ \dot{\xi}_{n-1} &= \xi_n \\ \dot{\xi}_n &= u \end{aligned}$$

where $n \geq 3$ and $f_1(\xi_n)$ is a C^∞ -smooth function such that all derivatives of $f_1(\xi_n)$ at $0 \in \mathbb{R}^n$ vanish but the function does not vanish identically in a neighborhood. Clearly, all the invariants $a^{[m]j,i+2}$ vanish but the system is not feedback linearizable because the distribution $\mathcal{D}^2 = \text{span}\{g, \text{ad}_f g\}$ is not involutive.

Now, we discuss the problem of feedback linearization of systems with uncontrollable linearization. We start with the following immediate generalization of Theorem 22.

PROPOSITION 9

A smooth (analytic) system Σ , is locally at ξ^0 , equivalent via a smooth (analytic) feedback to

$$\begin{aligned}\dot{x}_1 &= f^1(x_1) \\ \dot{x}_2 &= Ax_2 + Bv\end{aligned}$$

with (A, B) controllable (in Brunovsky form), where $(x_1, x_2) = (x_{1,1}, \dots, x_{1,r}, x_{2,1}, \dots, x_{2,n-r})$ if and only if it satisfies around ξ^0 the conditions (FL1) and (FL2) of Theorem 22. Moreover, $\dim x_2 = \dim \mathcal{D}^n(\xi^0) = n - r$.

In the single-input case we get the following.

COROLLARY 5

A smooth (analytic) single-input system Σ , is locally at ξ^0 , equivalent via a smooth (analytic) feedback to

$$\begin{aligned}\dot{x}_1 &= f^1(x_1) \\ \dot{x}_2 &= Ax_2 + Bv\end{aligned}$$

with (A, B) controllable, if and only if it satisfies in a neighborhood of ξ^0 :

$$(\text{FUL1})_{\text{SI}} \quad \dim \mathcal{D}^{n-r}(\xi) = \dim \mathcal{D}^{n-r+1}(\xi) = n - r$$

$$(\text{FUL2})_{\text{SI}} \quad \text{the distribution } \mathcal{D}^{n-r-1} \text{ is involutive}$$

If, additionally,

$$(\text{FUL3}) \quad \text{the eigenvalues of } J = (\partial f^1 / \partial x_1)(x_1^0) \text{ are nonresonant, where}$$

$$x^0 = (x_1^0, x_2^0) = (x_{1,1}^0, \dots, x_{1,r}^0, x_{2,1}^0, \dots, x_{2,n-r}^0),$$

then Σ^∞ (the infinite series expansion of Σ) is equivalent to the linear system

$$\begin{aligned}\dot{x}_1 &= Jx_1 \\ \dot{x}_2 &= Ax_2 + Bv\end{aligned}$$

via a formal feedback Γ^∞ .

Note that to check (FUL3), we do not need to bring the system into the partially linear form $\dot{x}_1 = f^1(x_1)$, $\dot{x}_2 = Ax_2 + Bv$, (which would, in general, require solving a system of first-order PDEs). Indeed, the condition (FUL1)_{SI} implies that

$$[f, \mathcal{D}^{n-r}] \subset \mathcal{D}^{n-r}$$

which in turn yields

$$[F\xi, \mathcal{D}^{n-r}] \subset \mathcal{D}^{n-r}$$

where $F\xi$ is the homogenous part of degree 1 of the vector field f and $\mathcal{D}^{n-r}$ is the homogenous part of degree 0 of the distribution $\mathcal{D}^{n-r}$. In other words, the linear map F leaves the linear subspace $\mathcal{D}^{n-r}$ of $\mathbb{R}^n$ invariant. It follows that F passes to the quotient, that is, defines the map $F_{\mathcal{D}^{n-r}} : \mathbb{R}^n / \sim \rightarrow \mathbb{R}^n / \sim$, where $x \sim \bar{x}$ if and only if $x - \bar{x} \in \mathcal{D}^{n-r}$. The eigenvalues of $\mathcal{J}$ are just the eigenvalues of $F_{\mathcal{D}^{n-r}}$. To calculate them, find a linear map $x = T\xi$ defining linear coordinates $(x_1, x_2) = (x_{1,1}, \dots, x_{1,r}, x_{2,1}, \dots, x_{2,n-r})$ such that $\mathcal{D}^{n-r} = \text{span}\{\partial/\partial x_{2,1}, \dots, \partial/\partial x_{2,n-r}\}$. Express $TFT^{-1}x = (\tilde{F}^1x, \tilde{F}^2x)$ then, clearly, $\tilde{F}^1x = Jx_1$ and the eigenvalues of J are just the eigenvalues of $F_{\mathcal{D}^{n-r}}$. Of course, the same analysis holds in the multi-input case.

We will end this section by giving a C^∞ -version of Corollary 5. Combining it with the linearizability results of Chen and Sternberg, we get the following:

COROLLARY 6

If a C^∞ -smooth system Σ satisfies the conditions (FUL1)_{SI}, (FUL2)_{SI}, (FUL3) of Corollary 5, then it is equivalent to the linear system

$$\begin{aligned}\dot{x}_1 &= Jx_1 \\ \dot{x}_2 &= Ax_2 + Bv\end{aligned}$$

via a C^∞ -smooth formal feedback Γ .

An analogous result in the analytic category is more restrictive and much more subtle and requires introducing the notion of Poincaré–Siegel domains [3].

4.10 Normal Forms for Discrete Time Control Systems

The method of normal forms has proved to be a useful approach in studying dynamical systems and control systems as illustrated throughout this chapter. The pioneer of this formal approach, Henri Poincaré, applied it to both continuous time dynamical systems (vector fields) and discrete time dynamical systems (maps) [67]. This chapter may sound incomplete if we omit to mention the work done for discrete time control systems. Normal forms for discrete time control systems have been studied using a similar approach to that presented in previous sections. Thus, quadratic and cubic

normal forms for discrete time control systems has been treated earlier [6, 30, 32, 61]. These normal forms have been utilized for stabilization of systems with uncontrollable linearization [27–29, 31, 33, 59–61]. Recently, a normal form, of any degree m , for discrete time control systems with controllable linearization was given by Hamzi and Tall [34]. We propose to explain briefly those results in this section.

The problem is to study the action of a feedback transformation

$$\Upsilon : \begin{array}{l} x = \phi(\xi) \\ u = \gamma(x, v) \end{array}$$

on a discrete time nonlinear control system

$$\Pi : \xi^+ = f(\xi, u), \quad \xi \in \mathbb{R}^n, \quad u \in \mathbb{R}$$

where $\xi^+ = \xi(k+1)$, and $f(\xi, u) = f(\xi(k), u(k))$ for any $k \in \mathbb{N}$. The transformation Υ brings Π to the system

$$\tilde{\Pi} : x^+ = \tilde{f}(x, v)$$

whose dynamics are given by

$$\tilde{f}(x, v) = \phi(f(\phi^{-1}(x), \gamma(x, v)))$$

We suppose that $(0, 0) \in \mathbb{R}^n \times \mathbb{R}$ is an equilibrium point, that is, $f(0, 0) = 0$, and we denote by

$$\Pi^{[1]} : \xi^+ = F\xi + Gu$$

its linearization at this point, where

$$F = \frac{\partial f}{\partial \xi}(0, 0), \quad G = \frac{\partial f}{\partial u}(0, 0)$$

We will assume that this linearization is controllable, that is

$$\text{span}\{F^i G : 0 \leq i \leq n-1\} = \mathbb{R}^n$$

Let us consider the Taylor series expansion Π^∞ of the system Π , given by

$$\Pi^\infty : \xi^+ = F\xi + Gu + \sum_{m=2}^{\infty} f^{[m]}(\xi, u) \quad (4.58)$$

and the Taylor series expansion Υ^∞ of the feedback transformation Υ , given by

$$\begin{aligned} \Upsilon^\infty : \quad x &= \phi(\xi) = T\xi + \sum_{m=2}^{\infty} \phi^{[m]}(\xi) \\ u &= \gamma(\xi, v) = K\xi + Lv + \sum_{m=2}^{\infty} \gamma^{[m]}(\xi, v) \end{aligned} \quad (4.59)$$

Throughout this section, and in particular in formula (4.58) and formula (4.59), the homogeneity of $f^{[m]}$ and $\gamma^{[m]}$ will be taken with respect to the variables $(\xi, u)^T$ and $(\xi, v)^T$, respectively.

We first notice that, because of the controllability assumption, there always exists a linear feedback transformation

$$\Upsilon^1 : \quad \begin{aligned} x &= T\xi \\ u &= K\xi + Lv \end{aligned}$$

bringing the linear part

$$\Pi^{[1]} : \xi^+ = F\xi + Gu$$

into the Brunovský canonical form

$$\Pi_{CF}^{[1]} : x^+ = Ax + Bv$$

Then we study, successively for $m \geq 2$, the action of the homogeneous feedback transformations

$$\Upsilon^m : \quad \begin{aligned} x &= \xi + \phi^{[m]}(\xi) \\ u &= v + \gamma^{[m]}(\xi, v) \end{aligned} \quad (4.60)$$

on the homogeneous systems

$$\Pi^{[m]} : \xi^+ = A\xi + Bu + f^{[m]}(\xi, u) \quad (4.61)$$

Consider another homogeneous system

$$\tilde{\Pi}^{[m]} : x^+ = Ax + Bv + \tilde{f}^{[m]}(x, v) \quad (4.62)$$

DEFINITION 2 We say that the homogeneous system $\Pi^{[m]}$, given by (4.61), is feedback equivalent to the homogeneous system $\tilde{\Pi}^{[m]}$, given by (4.62), if

there exist a homogeneous feedback transformation Υ^m , of the form (4.60), which brings the system $\Pi^{[m]}$ into the system $\tilde{\Pi}^{[m]}$ modulo terms in $P^{\geq m+1}$.

The following proposition is the analogue, for discrete time control systems, of Proposition 2, stated for continuous systems. It establishes the conditions of equivalence of two homogeneous systems.

PROPOSITION 10

The homogeneous feedback transformation Υ^m , defined by (4.60), brings the homogeneous system $\Pi^{[m]}$, given by (4.61), into the homogeneous system $\tilde{\Pi}^{[m]}$, given by (4.62), if and only if the following relations

$$\begin{aligned}\phi_j^{[m]}(A\xi + Bu) - \phi_{j+1}^{[m]}(\xi) &= \tilde{f}_j^{[m]}(\xi, u) - f_j^{[m]}(\xi, u) \\ \phi_n^{[m]}(A\xi + Bu) + \gamma^{[m]}(\xi) &= \tilde{f}_n^{[m]}(\xi, u) - f_n^{[m]}(\xi, u)\end{aligned}$$

hold for all $1 \leq j \leq n-1$.

The proof of this proposition is straightforward.

Main results. Let us denote the control by $v = x_{n+1}$, and for any $1 \leq i \leq n+1$,

$$\pi_i(x) = (x_1, \dots, x_i)$$

The main result for discrete time nonlinear control systems with controllable linearization is given in the following theorem.

THEOREM 23

The homogeneous control system $\Pi^{[m]}$, defined by (4.61), is feedback equivalent, by a homogeneous feedback transformation Υ^m of the form (4.60), to the normal form

$$\Pi_{\text{NF}}^{[m]} : x^+ = Ax + Bv + \tilde{f}^{[m]}(x, v)$$

where the components of the map $\tilde{f}^{[m]}(x, v)$ are given by

$$\tilde{f}_j^{[m]}(x, v) = \begin{cases} \sum_{i=j+2}^{n+1} x_1 x_i P_{j,i}^{[m-2]}(\pi_i(x)) & \text{if } 1 \leq j \leq n-1 \\ 0 & \text{if } j = n \end{cases} \quad (4.63)$$

We can notice the similarity of this normal form with that of continuous systems (4.8) with the notable difference that the polynomials $P_{j,i}^{[m-2]}(\pi_i(x))$, instead of being multiplied by x_i^2 , are multiplied by $x_1 x_i$. As the homogeneous feedback transformations Υ^m leave the terms of degree less than

m invariant, a successive application of Theorem 23 gives the following corollary.

COROLLARY 7

The control system Π^∞ , defined by (4.58), is feedback equivalent, by a formal feedback transformation Υ^∞ of the form (4.59), to the normal form

$$\Pi_{\text{NF}}^\infty : x^+ = Ax + Bv + \sum_{m=2}^{\infty} \bar{f}^{[m]}(x, v)$$

where for any $m \geq 2$, the components of the map $\bar{f}^{[m]}(x, v)$ are given by (4.63).

To illustrate our results, we consider the following example of a pendulum described elsewhere [83].

4.10.1 Example: Bressan and Rampazzo Pendulum

Consider the Bressan and Rampazzo pendulum [10, 83] described by the equations (see Example 7)

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= -g \sin \xi_3 + \xi_1 \xi_4^2 \\ \dot{\xi}_3 &= \xi_4 \\ \dot{\xi}_4 &= u\end{aligned}$$

where ξ_1 denotes the length of the pendulum, ξ_2 its velocity, ξ_3 the angle of the pendulum with respect to the horizontal, ξ_4 its angular velocity, and g the gravity constant.

We discretize the system by taking

$$\dot{\xi}_1 = \xi_1^+ - \xi_1, \quad \dot{\xi}_2 = \xi_2^+ - \xi_2, \quad \dot{\xi}_3 = \xi_3^+ - \xi_3, \quad \dot{\xi}_4 = \xi_4^+ - \xi_4$$

The system rewrites

$$\begin{aligned}\xi_1^+ &= \xi_1 + \xi_2 \\ \xi_2^+ &= \xi_2 - g \sin \xi_3 + \xi_1 \xi_4^2 \\ \xi_3^+ &= \xi_3 + \xi_4 \\ \xi_4^+ &= \xi_4 + u\end{aligned}$$

Consider the change of coordinates

$$z_1 = \xi_1$$

$$z_2 = \xi_2 + \xi_1$$

$$z_3 = -g \sin \xi_3 + 2\xi_2 + \xi_1$$

$$z_4 = -g \sin(\xi_4 + \xi_3) + 3\xi_2 - 2g \sin \xi_3 + 2\xi_1 \xi_4^2 + \xi_1$$

$$v = z_4^+$$

whose inverse is such that $\xi_4 = h(z_1, z_2, z_3, z_4)$ is a smooth function. This change of coordinates takes the system into the form

$$z_1^+ = z_2$$

$$z_2^+ = z_3 + z_1 h^2(z_1, z_2, z_3, z_4)$$

$$z_3^+ = z_4$$

$$z_4^+ = v$$

The function $h^2(z_1, z_2, z_3, z_4)$ could be decomposed as

$$h^2(z_1, z_2, z_3, z_4) = h_1(z_1, z_2, z_3) + z_4 h_2(z_1, z_2, z_3, z_4)$$

where the 1-jet at 0 of h_1 is zero and $h_2(0) = 0$. Put $H_1(z_1, z_2, z_3) = z_1 h_1(z_1, z_2, z_3)$.

The objective is to show that we can get rid of the terms $H_1(z_1, z_2, z_3)$. Let us suppose that the k -jet at 0 of H_1 is zero.

Consider the change of coordinates $\tilde{z}_1 = z_1$, $\tilde{z}_2 = z_2$, $\tilde{z}_3 = z_3 + H_1(z_1, z_2, z_3)$, $\tilde{z}_4 = \tilde{z}_3^+$. This change of coordinates, completed by the feedback $\tilde{z}_4^+ = w$, takes the system into the form

$$\tilde{z}_1^+ = \tilde{z}_2$$

$$\tilde{z}_2^+ = \tilde{z}_3 + \tilde{H}_1(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3) + \tilde{z}_1 \tilde{z}_4 \tilde{H}_2(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3, \tilde{z}_4)$$

$$\tilde{z}_3^+ = \tilde{z}_4$$

$$\tilde{z}_4^+ = w$$

where $\tilde{H}_1(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3)$ and $\tilde{H}_2(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3, \tilde{z}_4)$ are some smooth functions. It is enough to remark that the $(k+2)$ -jet at 0 of $\tilde{H}_1(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3)$ is zero because the 2-jet of $z_1 z_4 H_2(z)$ is zero. Then by iteration we can cancel terms $H_1(z_1, z_2, z_3)$

and put the system into the desired normal form

$$\begin{aligned}x_1^+ &= x_2 \\x_2^+ &= x_3 + x_1 x_4 H(x_1, x_2, x_3, x_4) \\x_3^+ &= x_4 \\x_4^+ &= v\end{aligned}$$

4.11 Symmetries of Control Systems

In this section, we will discuss relations between the canonical form $\Sigma_{\text{CF}}^\infty$ given in Section 4.4 and symmetries of nonlinear control systems.

Recently, there has been a growing interest in symmetries of nonlinear control systems. The structure of control systems possessing symmetries has been analyzed in several works [22, 24, 26]. The role of symmetries in the optimal control problems has been studied, among others, by Jurdjevic [47, 48] (for systems on Lie groups), van der Schaft [94], and Sussmann [93]. Jakubczyk [43] gave a complete characterization of symmetries in terms of symbols of control systems.

In this section, we study symmetries of single-input nonlinear control affine systems whose linear approximation, at an equilibrium point p , is controllable. We will discuss two results of the authors devoted, respectively, to stationary symmetries [73] and nonstationary symmetries [72]. The first, given in Section 4.11.2, states that “almost any” single-input control system, which is truly nonlinear (i.e., nonlinearizable via feedback) does not admit any stationary symmetry (i.e., any symmetry preserving the equilibrium point p). “Almost any” refers to all systems away from a small class of odd systems which admit one nontrivial stationary symmetry that is conjugated to minus identity by a diffeomorphism bringing the system to its canonical form $\Sigma_{\text{CF}}^\infty$ of Section 4.4. In Section 4.11.3, for the same class of systems and around an equilibrium point p , we study nonstationary symmetries (i.e., symmetries which do not preserve p). Our main result also states that for nonstationary symmetries a complete picture can be deduced from the canonical form. We prove that an analytic system, equivalent by an analytic feedback transformation to its canonical form, admits a nonstationary symmetry if and only if the drift vector field defining the canonical form is periodic with respect to the first variable and that a system admits a 1-parameter family of symmetries if and only if that drift vector field does not depend on the first variable. Moreover, we show that in the latter case the set of all symmetries is given either by exactly one 1-parameter family of symmetries (in the non-odd case) or by exactly

two 1-parameter families of symmetries (in the odd case). In the case when the feedback transformation, bringing the system to its canonical form, is given by a (not necessarily convergent) formal power series, we state in Section 4.11.4 that an analogous result holds for a formal infinitesimal symmetry. In fact, its existence is equivalent to the fact that the drift of the formal canonical form does not depend on the first variable.

We will also describe in Section 4.11.5 all symmetries of feedback linearizable systems [22, 24, 70] to show an enormous gap between the group of symmetries of feedback linearizable and nonlinearizable systems.

4.11.1 Symmetries

In this section, we will introduce the notion of symmetries of nonlinear control systems [26, 43, 73, 94]. Let us consider the system

$$\Pi : \dot{x} = F(x, u)$$

where $x \in X$, a smooth n -dimensional manifold and $u \in U$, a smooth m -dimensional manifold. The map $F : X \times U \rightarrow TX$ is assumed to be smooth with respect to (x, u) and for any value $u \in U$ of the control parameter, F defines a smooth vector field F_u on X , where $F_u(\cdot) = F(\cdot, u)$.

Consider the field of *admissible velocities* $\mathcal{F}$ associated to the system Π and defined as (see Section 4.1)

$$\mathcal{F}(x) = \{F_u(x) : u \in U\} \subset T_x X$$

We say that a diffeomorphism $\sigma : X \rightarrow X$ is a *symmetry* of Π if it preserves the field of admissible velocities $\mathcal{F}$, that is,

$$\sigma_* \mathcal{F} = \mathcal{F}$$

Recall that for any vector field f on X and any diffeomorphism $y = \phi(x)$ of X , we put

$$(\phi_* f)(y) = D\phi(\phi^{-1}(y)) \cdot f(\phi^{-1}(y))$$

A *local symmetry* at $p \in X$ is a local diffeomorphism σ of X_0 onto $\tilde{X}_0$, where X_0 and $\tilde{X}_0$ are, respectively, neighborhoods of p and $\sigma(p)$, such that

$$(\sigma_* \mathcal{F})(q) = \mathcal{F}(q)$$

for any $q \in \tilde{X}_0$.

A local symmetry σ at p is called a *stationary symmetry* if $\sigma(p) = p$ and a *nonstationary symmetry* if $\sigma(p) \neq p$.

Let us consider a single-input control affine system

$$\Sigma : \dot{x} = f(x) + g(x)u$$

where $x \in X, u \in U = \mathbb{R}$ and f and g are smooth vector fields on X . The field of admissible velocities for the system Σ is the following field of affine lines:

$$\mathcal{A}(x) = \{f(x) + ug(x) : u \in \mathbb{R}\} \subset T_x X$$

A specification of the aforementioned definition says that a diffeomorphism $\sigma : X \rightarrow X$ is a symmetry of Σ if it preserves the affine line field $\mathcal{A}$ (in other words, the affine distribution $\mathcal{A}$ of rank 1), that is, if

$$\sigma_* \mathcal{A} = \mathcal{A}$$

We will call $p \in X$ to be an *equilibrium point* of Σ if $0 \in \mathcal{A}(p)$. For any equilibrium point p there exists a unique $\tilde{u} \in \mathbb{R}$ such that $\tilde{f}(p) = 0$, where $\tilde{f}(p) = f(p) + \tilde{u}g(p)$. By the linear approximation of Σ at an equilibrium p we will mean the pair (F, G) , where $F = (\partial \tilde{f} / \partial x)(p)$ and $G = g(p)$.

We will say that Σ is an *odd system* at $p \in X$ if it admits a stationary symmetry at p , denoted by σ^- , such that

$$\frac{\partial \sigma^-}{\partial x}(p) = -Id$$

4.11.2 Symmetries of Single-Input Nonlinearizable Systems

In this section, we deal with single-input control affine systems of the form

$$\Sigma : \dot{x} = f(x) + g(x)u$$

where $x \in X$ and $u \in \mathbb{R}$. Our analysis will be local so we can assume that $X = \mathbb{R}^n$. Throughout this section, we will assume that the point p around which we work is an equilibrium, that is $f(p) = 0$ and, moreover, that $g(p) \neq 0$. We will prove that if Σ is not feedback linearizable (see Section 4.9), then the group of local symmetries of Σ around an equilibrium $p \in \mathbb{R}^n$ is very small. More precisely, the following result of the authors [72, 73] says that if Σ is analytic, then it admits at most two 1-parameter families of local symmetries. We will say that σ_c , where $c \in (-\varepsilon, \varepsilon) \subset \mathbb{R}$, is a *nontrivial 1-parameter analytic family of local symmetries* if each σ_c is a local analytic symmetry, $\sigma_{c_1} \neq \sigma_{c_2}$ if $c_1 \neq c_2$, and $\sigma_c(x)$ is jointly analytic with respect to (x, c) .

Assume that the system Σ is analytic. If the feedback transformation $\Gamma = (\phi, \alpha, \beta)$, bringing Σ^∞ into its canonical form Σ_{CF}^∞ , is analytic then

we will denote the analytic canonical form of Σ by Σ_{CF} . (i.e., the analytic system whose infinite Taylor expansion is given by $\Sigma_{\text{CF}}^\infty$).

THEOREM 24

Assume that the system Σ is analytic, the linear approximation (F, G) of Σ at an equilibrium point p is controllable and that Σ is not locally feedback linearizable at p . Assume, moreover, that the local feedback transformation, bringing Σ into its canonical form Σ_{CF} is analytic at p . Then there exists a local analytic diffeomorphism $\phi : X_0 \rightarrow \mathbb{R}^n$, where X_0 is a neighborhood of p , with the following properties:

1. *If σ is a local analytic stationary symmetry of Σ at p , then either $\sigma = \text{Id}$ or*

$$\phi \circ \sigma \circ \phi^{-1} = -\text{Id}$$

2. *If σ is a local analytic nonstationary symmetry of Σ at p , then*

$$\phi \circ \sigma \circ \phi^{-1} = T_c$$

where $c \in \mathbb{R}$ and T_c is either the translation $T_c = (x_1 + c, x_2, \dots, x_n)$ or T_c is replaced by $T_c^- = T_c \circ (-\text{Id}) = (-x_1 + c, -x_2, \dots, -x_n)$.

3. *If σ_c , $c \in (-\varepsilon, \varepsilon)$ is a nontrivial 1-parameter analytic family of local symmetries of Σ at p , then*

$$\phi \circ \sigma_c \circ \phi^{-1} = T_c$$

where T_c is as above, for $c \in (-\varepsilon, \varepsilon)$.

If we drop the assumption that Σ is equivalent to its canonical form Σ_{CF} by an analytic feedback transformation, then items (1) and (3) remain valid, with the local analytic diffeomorphisms ϕ being replaced by a formal diffeomorphism. This implies that the group of stationary symmetries of an analytic single-input control system is very small. Indeed, we have the following:

COROLLARY 8

Consider an analytic single-input control system Σ and assume that it is not feedback linearizable and its linear approximation at an equilibrium point p is controllable. Then Σ possesses at most two analytic stationary symmetries at p : an identity and, if Σ is odd, a symmetry σ^- satisfying $(\partial \sigma^- / \partial x)(p) = -\text{Id}$.

4.11.3 Symmetries of the Canonical Form

Symmetries take a very simple form if we bring the system into its canonical form. Indeed, we have the following result obtained by the authors ([72, 73] for proofs and details).

PROPOSITION 11

Assume that the system Σ is analytic, the linear approximation (F, G) of Σ at an equilibrium point p is controllable and Σ is not locally feedback linearizable at p . Assume, moreover, that the local feedback transformation, bringing Σ into its canonical form Σ_{CF} , is analytic at p :

1. Σ admits a nontrivial local stationary symmetry if and only if the drift $\bar{f}(x) = Ax + \sum_{m=0}^{\infty} \bar{f}^{[m]}(x)$ of the canonical form $\Sigma_{\text{CF}}^{\infty}$ satisfies

$$\bar{f}(x) = -\bar{f}(-x)$$

that is, the system is odd.

2. Σ admits a nontrivial local nonstationary symmetry if and only if the drift $\bar{f}(x)$ of the canonical form $\Sigma_{\text{CF}}^{\infty}$ satisfies

$$\bar{f}(x) = \bar{f}(T_c(x))$$

that is $\bar{f}$ is periodic with respect to x_1 .

3. Σ admits a nontrivial local 1-parameter family of symmetries if and only if the drift $\bar{f}(x)$ of the canonical form $\Sigma_{\text{CF}}^{\infty}$ satisfies

$$\bar{f}(x) = \bar{f}(x_2, \dots, x_n)$$

This result describes all symmetries around an equilibrium of any single-input nonlinear system that is not feedback linearizable and whose first-order approximation at the equilibrium is controllable. If we drop the assumption that Σ is equivalent to its canonical form Σ_{CF} by an analytic feedback transformation, then the “only if” statements in items (1) and (3) remain valid while in the “if” statements we have to replace local symmetries by formal symmetries, that is, by formal diffeomorphisms which preserve the field of admissible velocities [71], which we will do in the next section.

4.11.4 Formal Symmetries

We do not know whether, in general, the feedback transformation Γ^{∞} bringing the system to its canonical form $\Sigma_{\text{CF}}^{\infty}$ converges. If it does, Theorem 24 and Proposition 11 describe all local symmetries of Σ . If it does not,

the canonical form $\Sigma_{\text{CF}}^\infty$ is considered as a formal power series but even in this case it keeps, as we will show in the following, important information about symmetries.

We say that a vector field v on an open subset $X \subset \mathbb{R}^n$ is an *infinitesimal symmetry* of the system Σ if the (local) flow γ_t^v of v is a local symmetry of Σ , for any t for which it exists.

Consider the system Σ and denote by $\mathcal{G}$ the distribution spanned by the vector field g . We have the following characterization of infinitesimal symmetries.

PROPOSITION 12

A vector field v is an infinitesimal symmetry of Σ if and only if

$$[v, f] = 0 \bmod \mathcal{G}, \quad [v, g] = 0 \bmod \mathcal{G}$$

This characterization of infinitesimal symmetries justifies the following notion. We say that a vector field formal series

$$v^\infty(\xi) = \sum_{m=0}^{\infty} v^{[m]}(\xi)$$

is a *formal infinitesimal symmetry* of the system

$$\Sigma^\infty : \dot{\xi} = f(\xi) + g(\xi)u = \sum_{m=1}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}u \right)$$

if it satisfies

$$[v^\infty, f] = 0 \bmod \mathcal{G}, \quad [v^\infty, g] = 0 \bmod \mathcal{G}$$

Here, $[\cdot, \cdot]$ is understood as the Lie bracket of formal power series vector fields.

THEOREM 25

Consider the system Σ^∞ . Assume that its linear approximation (F, G) is controllable and that Σ^∞ is not feedback linearizable. The following conditions are equivalent:

1. Σ^∞ admits a formal infinitesimal symmetry.
2. The only formal infinitesimal symmetry of Σ^∞ is $v^\infty = (\phi^{-1})_*(\partial/\partial x_1)$, where ϕ is the diffeomorphism defining a feedback transformation Γ^∞ that brings Σ^∞ into its canonical form $\Sigma_{\text{CF}}^\infty$.

3. The canonical form $\Sigma_{\text{CF}}^\infty$ of Σ^∞ satisfies $\bar{f}^{[m]}(x) = \bar{f}^{[m]}(x_2, \dots, x_n)$, for any $m \geq m_0$, where the vector fields $\bar{f}^{[m]}$ are of the form (4.17)–(4.19).
4. For any $c_1 \in \mathbb{R}$, the translation $T_{c_1}(x) = (x_1 + c_1, x_2, \dots, x_n)^T$ is a symmetry of the canonical form $\Sigma_{\text{CF}}^\infty$.
5. The vector field $v_{\text{CF}}^\infty = \partial/\partial x_1$ is a formal infinitesimal symmetry of the canonical form $\Sigma_{\text{CF}}^\infty$.

This result, established in the formal category, provides the following necessary condition for the existence of analytic 1-parameter families of symmetries. Note that in the following we do not assume that the feedback transformation Γ^∞ , bringing Σ to its canonical form $\Sigma_{\text{CF}}^\infty$, converges.

PROPOSITION 13

Consider an analytic system Σ and assume that its linear approximation is controllable and the system is not feedback linearizable. If Σ admits a nontrivial analytic local 1-parameter group of symmetries σ_{c_1} , for $c_1 \in (-\varepsilon, \varepsilon)$, then the drift vector field of the canonical form $\Sigma_{\text{CF}}^\infty$ satisfies $\bar{f}^{[m]}(x) = \bar{f}^{[m]}(x_2, \dots, x_n)$, for any $m \geq m_0$.

We will end this section by giving a necessary condition for the existence of a family of local nonstationary symmetries which does not require to bring the system to its canonical form but only to normalize a finite number of terms.

Let m_0 denote the largest nonnegative integer such that for any $1 \leq k \leq n$, the distributions $\mathcal{D}^k = (g, \text{ad}_f g, \dots, \text{ad}_f^{k-1} g)$ have constant rank k and are involutive modulo terms of order $m_0 - 2$. It follows that the system Σ is feedback linearizable up to order $m_0 - 1$ [56]. We can thus bring the system to the form

$$\tilde{\Sigma} : \dot{x} = Ax + Bu + \bar{f}^{[m_0]}(x) + R(x, v) \quad (4.64)$$

where $R(x, v) \in V(x, v)^{\geq m_0+1}$ and (A, B) is in the Brunovský canonical form and the first nonlinearizable homogeneous vector field $\bar{f}^{[m_0]}$ whose components are given by

$$\bar{f}_j^{[m_0]}(x) = \begin{cases} \sum_{i=j+2}^n x_i^2 P_{j,i}^{[m_0-2]}(x_1, \dots, x_i) & \text{if } 1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \quad (4.65)$$

is in Kang normal form $\Sigma^{[m_0]}$.

PROPOSITION 14

Under the assumptions of Proposition 13, if $\bar{f}^{[m_0]}(x)$ depends on x_1 then Σ does not admit any nontrivial analytic local 1-parameter group of symmetries.

We would like to emphasize that the aforementioned condition is checkable via an algebraic calculation. In fact, bringing the terms of degree smaller than m_0 of Σ to their Kang normal form means simply to annihilate them (compare Section 4.9). In Section 4.3, we gave explicit polynomial transformations that bring a homogenous part of any degree of a system to Kang normal form $\Sigma^{[m_0]}$. Therefore, a successive use of those polynomial transformations, of degree 2 up to m_0 , brings Σ into $\tilde{\Sigma}$ for which we can apply Proposition 14.

4.11.5 Symmetries of Feedback Linearizable Systems

In the previous section, we proved that the group of symmetries of feedback nonlinearizable systems around an equilibrium is very small provided that the linear approximation at the equilibrium is controllable. A natural question is thus what are symmetries of feedback linearizable systems? In this section, we will show that symmetries of such systems form an infinite dimensional group parameterized by m arbitrary functions of m variables, where m is the number of controls. It is interesting to observe that just one nonlinearity, which is not removable by feedback, destroys this infinite dimensional group leaving, at most, two one-parameter families of symmetries (compare Theorem 24).

We will describe symmetries of linear systems in Brunovský canonical form and then of feedback linearizable systems. For the sake of simplicity, we will deal with systems with all controllability indices equal. Another description of symmetries of linear systems in Brunovský canonical form was given elsewhere [22, 24].

Consider a linear control system in the Brunovský canonical form with all controllability indices equal, say to n ,

$$\Lambda : \begin{aligned} \dot{x}^1 &= x^2 \\ &\vdots \\ \dot{x}^{n-1} &= x^n \\ \dot{x}^n &= v \end{aligned}$$

on $\mathbb{R}^N$, where $\dim v = m$, $\dim x^j = n$, $N = nm$.

Put $\pi^1(x) = x^1$. For any diffeomorphism μ of $\mathbb{R}^m$ we define $\mu^1 : \mathbb{R}^N \rightarrow \mathbb{R}^m$ by

$$\mu^1 = \mu \circ \pi^1$$

PROPOSITION 15

Consider the linear system Λ in Brunovský form

1. For any diffeomorphism μ of $\mathbb{R}^m$, the map

$$\lambda_\mu = \begin{pmatrix} \mu^1 \\ L_{Ax}\mu^1 \\ \vdots \\ L_{Ax}^{n-1}\mu^1 \end{pmatrix}$$

is a symmetry of Λ .

2. Conversely, if σ is a symmetry of Λ , then

$$\sigma = \lambda_\mu$$

for some diffeomorphism μ of $\mathbb{R}^m$.

Note that μ^1 is a map from $\mathbb{R}^N$ into $\mathbb{R}^m$ depending on the variables x^1 only. The transformation $\lambda_\mu : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is defined by successively differentiating this map with respect to the drift Ax . Item (1) claims that such a transformation is always a symmetry of the linear system Λ (in particular, a diffeomorphism) while item (2) claims that all symmetries of linear systems are always of this form.

REMARK 3

Clearly, an analogous result holds for local symmetries, that is, if μ is a local diffeomorphism of $\mathbb{R}^m$, then the corresponding λ_μ is a local symmetry of Λ and, conversely, any local symmetry of Λ is of the form λ_μ for some local diffeomorphism μ .

This local version of the earlier result will allow us to describe below all local symmetries of feedback linearizable systems.

Consider a control-affine system of the form

$$\Sigma : \dot{\xi} = f(\xi) + \sum_{i=1}^m g_i(\xi)u_i$$

where $\xi \in \Xi$, an N -dimensional manifold, and f and g_i for $1 \leq i \leq m$ are C^∞ vector fields on Ξ . We will say that Σ is feedback equivalent (or feedback linearizable) to a linear system of the form

$$\Lambda : \dot{x} = Ax + Bv$$

if there exists a feedback transformation of the form

$$\Gamma : \begin{aligned} x &= \Phi(\xi) \\ u &= \alpha(\xi) + \beta(\xi)v \end{aligned}$$

with $\beta(\xi)$ invertible, transforming Σ into Λ , compare Section 4.9. We say that Σ is locally feedback linearizable at ξ_0 if Φ is a local diffeomorphism at ξ_0 and α and β are defined locally around ξ_0 .

Define the following distributions:

$$\mathcal{D}^1 = \text{span} \{g_1, \dots, g_m\} \quad \text{and} \quad \mathcal{D}^{j+1} = \mathcal{D}^j + [f, \mathcal{D}^j].$$

The system Σ is, locally at ξ_0 , feedback equivalent to a linear system Λ , see Section 4.9, with all controllability indices equal to n , if and only if the distributions $\mathcal{D}^j$ are involutive and of constant rank jm for $1 \leq j \leq n$.

For any map $\varphi : \Xi_0 \rightarrow \mathbb{R}^m$, where Ξ_0 is a neighborhood of ξ_0 , put

$$\Phi_\varphi = \begin{pmatrix} \varphi \\ L_f \varphi \\ \vdots \\ L_f^{n-1} \varphi \end{pmatrix}$$

Note that Φ_φ is a map from Ξ_0 in $\mathbb{R}^N$. If the map $\varphi = (\varphi_1, \dots, \varphi_m)$ is chosen such that

$$(\mathcal{D}^{n-1})^\perp = \text{span} \{d\varphi\} = \text{span} \{d\varphi_1, \dots, d\varphi_m\}$$

then it is well known [38, 66] that Φ_φ is a local diffeomorphism of an open neighborhood Ξ_φ of ξ_0 onto $X_\varphi = \Phi_\varphi(\Xi_\varphi)$, an open neighborhood of $x_0 = \Phi_\varphi(\xi_0)$, and gives local linearizing coordinates for Σ in Ξ_φ . To keep the notation coherent, we will denote by ξ , with various indices, points of Ξ_φ , by x , with various indices, points of $X_\varphi = \Phi_\varphi(\Xi_\varphi) \subset \mathbb{R}^N$, and by y , with various indices, points of $\pi^1(X_\varphi) \subset \mathbb{R}^m$, where π^1 is the projection $\pi^1(x) = x^1$.

Combining this result with Proposition 15, we get the following complete description of local symmetries of feedback linearizable systems with equal controllability indices. The notation $\text{Diff}(\mathbb{R}^m; y_0, \tilde{y}_0)$ will stand for the family of all local diffeomorphisms of $\mathbb{R}^m$ at y_0 transforming y_0 into $\tilde{y}_0$

(more precisely, all diffeomorphisms germs with the base point y_0 and its image $\tilde{y}_0$).

THEOREM 26

Let the system Σ be locally feedback linearizable at ξ_0 with equal controllability indices. Fix $\varphi : \Xi_0 \rightarrow \mathbb{R}^m$ such that $(\mathcal{D}^{n-1})^\perp = \text{span } \{d\varphi\} = \text{span } \{d\varphi_1, \dots, d\varphi_m\}$:

1. Let $\mu \in \text{Diff}(\mathbb{R}^m; y_0, \tilde{y}_0)$, where $y_0 = \pi^1(x_0)$ and $\tilde{y}_0 = \pi^1(\lambda_\mu(x_0))$, such that $\lambda_\mu(x_0) \in X_\varphi$. Then

$$\sigma_{\mu,\varphi} = \Phi_\varphi^{-1} \circ \lambda_\mu \circ \Phi_\varphi$$

is a local symmetry of Σ at ξ_0 .

2. Conversely, if σ is a local symmetry of Σ at ξ_0 , such that $\sigma(\xi_0) \in \Xi_\varphi$, then there exists $\mu \in \text{Diff}(\mathbb{R}^m; y_0, \tilde{y}_0)$, where $y_0 = \pi^1(x_0)$, $\tilde{y}_0 = \pi^1(\tilde{x}_0)$, $\tilde{x}_0 = \Phi_\varphi(\sigma(\xi_0))$ such that

$$\sigma = \sigma_{\mu,\varphi}$$

Moreover, $\sigma_{\mu,\varphi} = \Phi_\varphi^{-1} \circ \lambda_\mu \circ \Phi_\varphi = \Phi_\varphi^{-1} \circ \Phi_{\mu \circ \varphi}$.

The structure of symmetries of feedback linearizable systems is thus summarized by the following diagram.

$$\begin{array}{ccc} (\Sigma, \xi_0) & \xrightarrow{\sigma_{\mu,\varphi}} & (\Sigma, \tilde{\xi}_0) \\ \Phi_\varphi \downarrow & \searrow \Phi_{\mu \circ \varphi} & \downarrow \Phi_\varphi \\ (\Lambda, x_0) & \xrightarrow{\lambda_\mu} & (\Lambda, \tilde{x}_0) \end{array}$$

Item (1) states that composing a linearizing transformation Φ_φ with a symmetry λ_μ of the linear equivalent Λ of Σ and with the inverse Φ_φ^{-1} we get a symmetry of Σ , provided that the image $\tilde{x}_0 = \lambda_\mu(x_0)$ belongs to X_φ (otherwise the composition is not defined). Item (2) asserts that any local symmetry of a feedback linearizable system is of this form. Moreover, any local symmetry can be expressed as a composition of one linearizing transformation with the inverse of another linearizing transformation. Indeed, observe that for any fixed φ , the map $\Phi_{\mu \circ \varphi}$, for $\mu \in \text{Diff}(\mathbb{R}^m; y_0, \tilde{y}_0)$, gives a linearizing diffeomorphism, and taking all $\mu \in \text{Diff}(\mathbb{R}^m; y_0, \tilde{y}_0)$

for all $\tilde{y}_0 \in \pi^1(X_\varphi)$, the corresponding maps $\Phi_{\mu \circ \varphi}$ provide all linearizing transformations around ξ_0 .

It follows from item (2) that the group of symmetries of feedback linearizable systems is infinite dimensional and parameterized by m functions of m variables. It is interesting to observe that just one nonlinearity, which is not removable by feedback, destroys this infinite dimensional group leaving, at most, two one-parameter families of symmetries (compare Theorem 24).

4.12 Feedforward and Strict Feedforward Forms

In this section, we study the problem of transforming a single-input nonlinear control system to feedforward form and strict feedforward form via a static state feedback. We provide checkable necessary and sufficient conditions (which involve the homogeneous m -invariants defined in Section 4.3) to bring the homogeneous terms of any fixed degree of the system into homogeneous feedforward form. If those conditions are satisfied, this leads to a constructive procedure which transforms the system, step by step, into feedforward or strict feedforward form. We illustrate our solution by analyzing the four-dimensional case. In particular, we compute the codimension of four-dimensional systems that are feedback equivalent to the feedforward form and strict feedforward form.

This section is organized as follows. In Section 4.12.1, we will define the class of feedforward and strict feedforward systems, in both general and control-affine cases. We will also fix some notations used throughout the whole section. In Section 4.12.2, we will introduce feedforward and strict feedforward normal forms. Then we will present a step-by-step method transforming a given system to the feedforward or strict feedforward form (whenever it is possible): in Section 4.12.3 for the first nonlinearizable term and in Section 4.12.4 for terms of an arbitrary degree. We will illustrate our approach by analyzing feedforward and strict feedforward systems on $\mathbb{R}^4$ in Section 4.12.5. Finally, we will discuss the geometry of feedforward and strict feedforward systems in Section 4.12.6 and their symmetries in Section 4.12.7.

4.12.1 Introduction and Notations

Consider a single-input nonlinear control system of the form

$$\Pi : \dot{\xi} = F(\xi, u)$$

where $\xi \in \mathbb{R}^n$ and $u \in \mathbb{R}$. We say that the system Π is in *feedforward form* (resp. in *strict feedforward form*) if we have

$$\begin{array}{ll} \dot{\xi}_1 = F_1(\xi_1, \dots, \xi_n, u) & \dot{\xi}_1 = F_1(\xi_2, \dots, \xi_n, u) \\ \dot{\xi}_2 = F_2(\xi_2, \dots, \xi_n, u) & \dot{\xi}_2 = F_2(\xi_3, \dots, \xi_n, u) \\ \vdots & \left(\text{resp.} \quad \vdots \right) \\ \dot{\xi}_{n-1} = F_{n-1}(\xi_{n-1}, \xi_n, u) & \dot{\xi}_{n-1} = F_{n-1}(\xi_n, u) \\ \dot{\xi}_n = F_n(\xi_n, u) & \dot{\xi}_n = F_n(u) \end{array}$$

One of the most appealing features of the system in (strict) feedforward form is that we can construct a stabilizing feedback for them. This important result goes back to Teel [90] and has been followed by a growing literature on stabilization and tracking for systems in (strict) feedforward form [5, 39, 64, 65, 76, 91].

Feedforward systems can be viewed as duals of feedback linearizable systems. To see this, recall that in the single-input case, the class of feedback linearizable systems coincides with that of flat systems. Single-input flat systems are defined as systems for which we can find a function of the state that, together with its *derivatives*, gives all the states and the control of the system [19, 20, 41, 68]. In a dual way, for systems in strict feedforward form, we can find all states via a successive *integration* starting from a function of the control. Indeed, knowing $u(t)$, we integrate $F_n(u(t))$ to get $\xi_n(t)$; then we integrate $F_{n-1}(\xi_n(t), u(t))$ to get $\xi_{n-1}(t)$; we continue doing this, and finally we integrate $F_1(\xi_2(t), \dots, \xi_n(t), u(t))$ to get $\xi_1(t)$. For feedforward systems, solutions can be found by solving scalar differential equations: for each component we have to solve one scalar differential equation.

It is therefore natural to ask which systems are equivalent to one of the just defined feedforward forms. In Ref. [63], the problem of transforming a system, linear with respect to controls, into (strict) feedforward form via a diffeomorphism (i.e., via a nonlinear change of coordinates), was studied. A geometric description of systems transformable into feedforward form, either via a diffeomorphism or via feedback, has been given elsewhere [4]. Similar conditions for the strict feedforward form have recently been proposed by the authors [74], where relations between strict feedforward systems and the notion of symmetries (as defined in Section 4.11) are studied, see Section 4.12.7. The conditions described in Refs. [4, 74] (which we recall in Section 4.12.6) are intrinsic and explain the geometry of the problem but in most cases are not checkable. In Refs. [80, 82, 84], we proposed a constructive procedure which allows to verify, step by step, whether a given system is feedback equivalent to the feedforward or strict feedforward form and to bring it to that form whenever it is possible.

Our solutions were inspired by and are based on a formal approach to the feedback equivalence problem described in Section 4.3 and thus constitute a good example of the strength of the formal approach.

We will be dealing with control-affine systems of the form

$$\Sigma : \dot{\xi} = f(\xi) + g(\xi)u$$

where $\xi \in \mathbb{R}^n$ and $u \in \mathbb{R}$. As usual, we assume that $f(0) = 0$ and $g(0) \neq 0$. A specification of the general definition to the control-affine case implies that Σ is in *feedforward form*, or that it is a *feedforward system*, if we have

$$f(\xi) = \begin{pmatrix} f_1(\xi_1, \dots, \xi_n) \\ f_2(\xi_2, \dots, \xi_n) \\ \vdots \\ f_{n-1}(\xi_{n-1}, \xi_n) \\ f_n(\xi_n) \end{pmatrix} \quad \text{and} \quad g(\xi) = \begin{pmatrix} g_1(\xi_1, \dots, \xi_n) \\ g_2(\xi_2, \dots, \xi_n) \\ \vdots \\ g_{n-1}(\xi_{n-1}, \xi_n) \\ g_n(\xi_n) \end{pmatrix}$$

Similarly, Σ is in *strict feedforward form* (equivalently, it is a *strict feedforward system*), if we have

$$f(\xi) = \begin{pmatrix} f_1(\xi_2, \dots, \xi_n) \\ f_2(\xi_3, \dots, \xi_n) \\ \vdots \\ f_{n-1}(\xi_n) \\ f_n \end{pmatrix} \quad \text{and} \quad g(\xi) = \begin{pmatrix} g_1(\xi_2, \dots, \xi_n) \\ g_2(\xi_3, \dots, \xi_n) \\ \vdots \\ g_{n-1}(\xi_n) \\ g_n \end{pmatrix}$$

where the components f_n and g_n are constant and satisfy $f_n = 0$ (because 0 is assumed to be an equilibrium) and $g_n \neq 0$.

To present our step-by-step approach, together with the system

$$\Sigma : \dot{\xi} = f(\xi) + g(\xi)u$$

we will consider its infinite Taylor series expansion

$$\Sigma^\infty : \dot{\xi} = F\xi + Gu + \sum_{m=2}^{\infty} (f^{[m]}(\xi) + g^{[m-1]}(\xi)u)$$

where $F = (\partial f / \partial \xi)(0)$ and $G = g(0)$.

Consider the infinite Taylor series expansion Γ^∞ of the feedback transformation Γ given by

$$\Gamma^\infty : \begin{aligned} x &= \phi(\xi) = T\xi + \sum_{m=2}^{\infty} \phi^{[m]}(\xi) \\ u &= \alpha(\xi) + \beta(\xi)v = K\xi + Lv + \sum_{m=2}^{\infty} \left(\alpha^{[m]}(\xi) + \beta^{[m-1]}(\xi)v \right) \end{aligned}$$

where T is an invertible matrix and $L \neq 0$. We will try (whenever possible) to bring the system Σ^∞ into the (strict) feedforward form step by step by analyzing the action of Γ^∞ .

Notations. We will use normal forms and transformations which have been already defined in the paper but we will also introduce normal forms specific for this section. The symbols $\Sigma^{[m]}$, $\Sigma^{[\leq m]}$, and Σ^∞ will stand for the systems under consideration: homogenous, polynomial, and formal, respectively. Their state vector will be denoted by ξ and their control by u . The system $\Sigma^{[m]}$ (resp. $\Sigma^{[\leq m]}$ and Σ^∞) transformed via a feedback transformation Γ^m (resp. $\Gamma^{\leq m}$ and Γ^∞) will be denoted by $\tilde{\Sigma}^{[m]}$ (resp. $\tilde{\Sigma}^{[\leq m]}$ and $\tilde{\Sigma}^\infty$). Its state vector will be denoted by x , its control by v , and the vector fields, defining its dynamics, by $\tilde{f}^{[m]}$ and $\tilde{g}^{[m-1]}$. Feedback equivalence of systems $\Sigma^{[m]}$ and $\tilde{\Sigma}^{[m]}$ and of systems $\Sigma^{[\leq m]}$ and $\tilde{\Sigma}^{[\leq m]}$ will be established via a smooth feedback, through a homogeneous feedback Γ^m in the former case and through a polynomial feedback $\Gamma^{\leq m}$ in the latter. On the other hand, feedback equivalence of systems Σ^∞ and $\tilde{\Sigma}^\infty$ will be established via a formal feedback Γ^∞ .

We will use three kinds of normal forms for systems: Kang normal forms, feedforward normal forms, and strict feedforward normal forms. The symbol “bar” will correspond to the vector field $\tilde{f}^{[m]}$ defining Kang normal forms $\Sigma_{\text{NF}}^{[m]}$, $\Sigma_{\text{NF}}^{[\leq m]}$, and $\Sigma_{\text{NF}}^\infty$. The symbol “hat” will correspond to the vector field $\hat{f}^{[m]}$ defining feedforward normal forms $\Sigma_{\text{FNF}}^{[m]}$, $\Sigma_{\text{FNF}}^{[\leq m]}$, and $\Sigma_{\text{FNF}}^\infty$ and strict feedforward normal forms $\Sigma_{\text{SFNF}}^{[m]}$, $\Sigma_{\text{SFNF}}^{[\leq m]}$, and $\Sigma_{\text{SFNF}}^\infty$. Analogously, the m -invariants of the system $\Sigma^{[m]}$ will be denoted by $a^{[m]j,i+2}$, the m -invariants of the system $\Sigma_{\text{NF}}^{[m]}$ by $\bar{a}^{[m]j,i+2}$, and the m -invariants of the systems $\Sigma_{\text{FNF}}^{[m]}$ or $\Sigma_{\text{SFNF}}^{[m]}$ by $\hat{a}^{[m]j,i+2}$.

4.12.2 Feedforward and Strict Feedforward Normal Forms

We assume throughout this section that the linear part (F, G) of the system Σ^∞ , given by (4.66), is controllable, and thus we can assume, without loss

of generality, that the system is in the form

$$\Sigma^\infty : \dot{\xi} = A\xi + Bu + \sum_{m=2}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right) \quad (4.66)$$

where (A, B) is in the Brunovský canonical form.

Recall from Section 4.3, that (as proved by Kang [50], compare also Ref. [83]) any nonlinear system of the form (4.66) can be put, via a formal feedback transformation Γ^∞ , to the following Kang normal form:

$$\Sigma_{\text{NF}}^\infty : \dot{x} = Ax + Bv + \sum_{m=2}^{\infty} \bar{f}^{[m]}(x)$$

where for any $m \geq 2$,

$$\bar{f}_j^{[m]}(x) = \begin{cases} \sum_{i=j+2}^n x_i^2 P_{j,i}^{[m-2]}(x_1, \dots, x_i) & \text{if } 1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \quad (4.67)$$

It is natural to ask whether it is possible to bring a system that is feedback equivalent to feedforward form (resp. strict feedforward form) to the Kang normal form (4.67) which would be simultaneously feedforward (resp. strict feedforward), that is, which would satisfy

$$P_{j,i}^{[m-2]}(x) = P_{j,i}^{[m-2]}(x_j, \dots, x_i) \quad \left(\text{resp. } P_{j,i}^{[m-2]}(x) = P_{j,i}^{[m-2]}(x_{j+1}, \dots, x_i) \right)$$

Although, this is always possible for the first nonlinearizable term (see Theorem 28), the answer to the question is, in general, negative. For this reason we will introduce the following notions.

DEFINITION 3 *Strict feedforward normal form* is the system

$$\Sigma_{\text{SFNF}}^\infty : \dot{x} = Ax + Bv + \sum_{m=2}^{\infty} \hat{f}^{[m]}(x)$$

such that for any $m \geq 2$,

$$\hat{f}_j^{[m]}(x) = \begin{cases} c_{m,j} x_{j+1}^m + \sum_{i=j+2}^n x_i^2 P_{j,i}^{[m-2]}(x_{j+1}, \dots, x_i) & 1 \leq j \leq n-2 \\ 0 & n-1 \leq j \leq n \end{cases}$$

where $c_{m,j} \in \mathbb{R}$ and $P_{j,i}^{[m-2]}$ are homogeneous polynomials of degree $m-2$, depending on the indicated variables.

In the feedforward case we introduce similarly:

DEFINITION 4 *Feedforward normal form* is the system

$$\Sigma_{\text{FNF}}^{\infty} : \dot{x} = Ax + Bv + \sum_{m=2}^{\infty} \hat{f}^{[m]}(x)$$

such that for any $m \geq 2$,

$$\hat{f}_j^{[m]}(x) = \begin{cases} x_j k_j^{[m-1]}(x_j, x_{j+1}) + \sum_{i=j+2}^n x_i^2 P_{j,i}^{[m-2]}(x_j, \dots, x_i) & 1 \leq j \leq n-2 \\ 0 & n-1 \leq j \leq n \end{cases}$$

where $k_j^{[m-1]}$ and $P_{j,i}^{[m-2]}$ are homogeneous polynomials, of degree $m-1$ and $m-2$, respectively, depending on the indicated variables.

Usefulness of feedforward and strict feedforward normal forms is justified by the following theorem.

THEOREM 27

The system Σ^{∞} , given by (4.66), is feedback equivalent to the feedforward form (resp. strict feedforward form) if and only if it is feedback equivalent to the feedforward normal form $\Sigma_{\text{FNF}}^{\infty}$ (resp. strict feedforward normal form $\Sigma_{\text{SFNF}}^{\infty}$).

4.12.3 Feedforward and Strict Feedforward Form: First Nonlinearizable Term

Consider the system Σ^{∞} , given by (4.66). Our goal is to study, whether it is possible to bring Σ^{∞} to the feedforward (resp. strict feedforward form) and, if possible, to do it step by step. Assume that Σ^{∞} is feedback linearizable up to order $m_0 - 1$. As proved by Krener [56], m_0 is the largest integer such that all distributions

$$\mathcal{D}^k = \text{span}\{g, \text{ad}_f g, \dots, \text{ad}_f^{k-1} g\}$$

for $1 \leq k \leq n-1$, are involutive modulo terms of order $m_0 - 2$. Since Σ^{∞} is feedback linearizable up to order $m_0 - 1$, it is also feedback equivalent to the strict feedforward (in particular, to the feedforward) form up to the same order, and we can thus assume without loss of generality that the system Σ^{∞} is in feedforward (resp. strict feedforward) normal form up to order $m_0 - 1$ (see Theorem 27), that is, it takes the form

$$\Sigma^{[\leq m_0]} : \dot{\xi} = A\xi + Bu + \sum_{m=2}^{m_0-1} h^{[m]}(\xi) + f^{[m_0]}(\xi) + g^{[m_0-1]}(\xi)u \quad (4.68)$$

modulo terms in $V^{\geq m_0+1}(\xi, u)$, where for any $2 \leq m \leq m_0 - 1$ we have

$$h_j^{[m]}(\xi) = \begin{cases} \xi_j k_j^{[m-1]}(\xi_j, \xi_{j+1}) & \text{if } 1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \quad (4.69)$$

$$\left(\text{resp. } h_j^{[m]}(\xi) = \begin{cases} c_{m,j} \xi_{j+1}^m & \text{if } 1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \right) \quad (4.70)$$

Let us denote by $a^{[m_0]j,i+2}$ the m_0 -invariants associated to the homogeneous part of degree m_0 of the system (4.68) and (4.69) or (4.68)–(4.70). Recall that

$$\Delta = \{(j, i) \in \mathbb{N} \times \mathbb{N} : 1 \leq j \leq n-2, 0 \leq i \leq n-j-2\}$$

We have the following result concerning the first nonlinearizable term:

THEOREM 28

Consider the system $\Sigma^{[\leq m_0]}$, given by (4.68):

1. There exists a transformation $\Gamma^{\leq m_0}$ bringing the system (4.68) into the feedforward form, up to order m_0 if and only if

$$L_{A^{n-q}B} a^{[m_0]j,i+2} = 0 \quad (4.71)$$

for any $(j, i) \in \Delta$ and any $1 \leq q \leq j-1$.

2. There exists a transformation $\Gamma^{\leq m_0}$ bringing the system (4.68) into the strict feedforward form, up to order m_0 if and only if

$$L_{A^{n-q}B} a^{[m_0]j,i+2} = 0 \quad (4.72)$$

for any $(j, i) \in \Delta$ and any $1 \leq q \leq j$.

COROLLARY 9

If there exists a transformation Γ^∞ bringing the system

$$\Sigma^\infty : \dot{\xi} = A\xi + Bu + \sum_{m=m_0}^{\infty} \left(f^{[m]}(\xi) + g^{[m-1]}(\xi)u \right)$$

to feedforward form (resp. strict feedforward form), then the condition (4.71) (resp. (4.72)) is satisfied for any $(j, i) \in \Delta$ and any $1 \leq q \leq j-1$ (resp. $1 \leq q \leq j$).

In other words, the aforementioned result says that if a system is feedback equivalent to the feedforward form (resp. strict feedforward form), then, after having linearized lower order terms, the first nonlinearizable term must be feedforward (resp. strict feedforward) when transformed to the Kang normal form. This is the case if and only if the condition (4.71) (resp. (4.72)) is satisfied. If $m_0 = 2$, then the 2-invariants $a^{[2]j,i+2}$ are constant and the conditions (4.71) and (4.72) are automatically satisfied, which implies that any system is equivalent to the strict feedforward form (in particular, to the feedforward form) up to order 2. Actually, in this case, the Kang–Krener normal form $\Sigma_{\text{NF}}^{[2]}$ (recalled just following Theorem 3) is strict feedforward and can serve as a strict feedforward normal form (we do not have to add the vector field $h^{[2]}$).

COROLLARY 10 (Kang–Krener)

If $m_0 = 2$ then the system (4.68) is always equivalent to the strict feedforward form (in particular, to the feedforward form) up to order 2.

4.12.4 Feedforward and Strict Feedforward Forms: The General Step

According to Theorem 28, the Kang normal form of the first nonlinearizable term of a system, which is feedback equivalent to feedforward (resp. strict feedforward form), must be feedforward (resp. strict feedforward). We will see in the following text, that the situation gets different when we proceed to higher-order terms. Let us assume that the system Σ^∞ , given by (4.66), is in feedforward (resp. strict feedforward) normal form up to order $m_0 + l - 1$, that is, Σ^∞ takes the form

$$\begin{aligned} \Sigma^\infty : \dot{\xi} = A\xi + Bu + \sum_{m=2}^{m_0+l-1} h^{[m]}(\xi) \\ + \sum_{m=m_0}^{m_0+l-1} \bar{f}^{[m]}(\xi) + f^{[m_0+l]}(\xi) + g^{[m_0+l-1]}(\xi)u + R(\xi, u) \end{aligned} \quad (4.73)$$

where $R(\xi, u) \in V^{\geq m_0+l+1}(\xi, u)$ and for any $m_0 \leq m \leq m_0 + l - 1$,

$$\bar{f}_j^{[m]}(\xi) = \begin{cases} \sum_{i=j+2}^n \xi_i^2 P_{j,i}^{[m-2]}(\xi_j, \dots, \xi_i) & \text{if } 1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \quad (4.74)$$

and

$$h_j^{[m]}(\xi) = \begin{cases} \xi_j k_j^{[m-1]}(\xi_j, \xi_{j+1}) & \text{if } 1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \quad (4.75)$$

$$\left(\text{resp. } h_j^{[m]}(\xi) = \begin{cases} c_{m,j} \xi_{j+1}^m & \text{if } 1 \leq j \leq n-2 \\ 0 & \text{if } n-1 \leq j \leq n \end{cases} \right) \quad (4.76)$$

From the definition of m_0 , it follows that there exists $1 \leq j \leq n-2$ such that $\bar{f}_j^{[m_0]} \neq 0$. Throughout we will assume that $\bar{f}_{n-2}^{[m_0]} \neq 0$, which simplifies the exposition. The analysis of the general case, although more technical, follows the same line (see Ref. [80] for the strict feedforward form).

Crucial objects in studying the strict feedforward case are $a_{k,0}^{[m_0+l]j,i+2}$, which are the homogeneous invariants associated to the homogeneous system

$$\Sigma_{k,0}^{[m_0+l]} : \dot{\xi} = A\xi + Bu + [\bar{f}^{[m_0]}, Y_k^{[l+1]}](\xi)$$

where

$$Y_k^{[l+1]} = Y_{k,0}^{[l+1]} = \xi_k^{l+1} \frac{\partial}{\partial \xi_k} + L_{A\xi} \xi_k^{l+1} \frac{\partial}{\partial \xi_{k+1}} + \cdots + L_{A\xi}^{n-k} \xi_k^{l+1} \frac{\partial}{\partial \xi_n}$$

For any $2 \leq k \leq n-2$ and any $0 \leq s \leq l+1$, consider the homogeneous vector fields

$$\begin{aligned} Y_{k,s}^{[l+1]} &= \xi_{k-1}^s \xi_k^{l+1-s} \frac{\partial}{\partial \xi_k} + L_{A\xi} \left(\xi_{k-1}^s \xi_k^{l+1-s} \right) \frac{\partial}{\partial \xi_{k+1}} \\ &+ \cdots + L_{A\xi}^{n-k-1} \left(\xi_{k-1}^s \xi_k^{l+1-s} \right) \frac{\partial}{\partial \xi_n} \end{aligned}$$

In studying the feedforward case, crucial objects are $a^{[m_0+l]j,i+2}$, which are the $(m_0 + l)$ -invariants of the homogeneous system $\Sigma^{[m_0+l]}$, defined by (4.73)–(4.75), and $a_{k,s}^{[m_0+l]j,i+2}$, the $(m_0 + l)$ -invariants of the homogeneous system

$$\Sigma_{k,s}^{[m_0+l]} : \dot{\xi} = A\xi + Bu + [\bar{f}^{[m_0]}, Y_{k,s}^{[l+1]}](\xi)$$

Our two main results of this section can be stated as follows. In the strict feedforward case we have (see [80, 84] for the proof).

THEOREM 29

The system Σ^∞ , defined by (4.73), (4.74), (4.76), is feedback equivalent to the strict feedforward form, up to order $m_0 + l$, if and only if there exist real constants $\sigma_{2,0}, \sigma_{3,0}, \dots, \sigma_{n-2,0}$ such that for any $(j, i) \in \Delta$ and any $1 \leq q \leq j$

$$L_{A^{n-q}B} \left(a^{[m_0+l]j,i+2} - \sum_{k=2}^{n-2} \sigma_{k,0} a_{k,0}^{[m_0+l]j,i+2} \right) = 0 \quad (4.77)$$

In the feedforward case we have (see [82] for the proof and comments).

THEOREM 30

The system Σ^∞ , defined by (4.73–4.75), is feedback equivalent, up to order $m_0 + l$, to the feedforward form if and only if there exist real constants $\sigma_{k,s}$ for $2 \leq k \leq n - 2$ and $1 \leq s \leq l + 1$ such that for any $(j, i) \in \Delta$ and any $1 \leq q \leq j - 1$

$$L_{A^{n-q}B} \left(a^{[m_0+l]j,i+2} - \sum_{k=2}^{n-2} \sum_{s=1}^{l+1} \sigma_{k,s} a_{k,s}^{[m_0+l]j,i+2} \right) = 0 \quad (4.78)$$

Notice that (4.77) is an invariant way of expressing the fact that

$$a^{[m_0+l]j,i+2} - \sum_{k=2}^{n-2} \sigma_{k,0} a_{k,0}^{[m_0+l]j,i+2} = Q_{j,i}^{[m_0+l-2]}$$

where $Q_{j,i}^{[m_0+l-2]} = Q_{j,i}^{[m_0+l-2]}(\xi_{j+1}, \dots, \xi_i)$ are homogeneous polynomials of degree $m_0 + l - 2$ depending on the indicated variables only. Similarly, (4.78) is an invariant way of expressing the fact that

$$a^{[m_0+l]j,i+2} - \sum_{k=2}^{n-2} \sum_{s=1}^{l+1} \sigma_{k,s} a_{k,s}^{[m_0+l]j,i+2} = Q_{j,i}^{[m_0+l-2]}$$

where $Q_{j,i}^{[m_0+l-2]} = Q_{j,i}^{[m_0+l-2]}(\xi_j, \dots, \xi_i)$ are homogeneous polynomials of degree $m_0 + l - 2$ depending on the indicated variables only.

Observe that checking the conditions (4.72) and (4.77) or the conditions (4.71) and (4.78) involves only differentiation of polynomials and algebraic operations. Therefore Theorem 28 (1) followed by a successive application of Theorem 30 (resp. Theorem 28 (2) followed by a successive application of Theorem 29) yields to a constructive procedure that allows us to check whether a given system can be transformed into feedforward (resp. strict feedforward) form. Moreover, for any system satisfying the conditions (4.71) and (4.78) (resp. (4.72) and (4.77)), we can calculate, step by step, the explicit feedback transformations bringing it into feedforward (resp. strict feedforward) form using transformations constructed in Refs. [79, 83] and presented in Section 4.3.

Finally, observe that the condition (4.78) can be seen as a natural generalization of (4.77). Indeed, in the second sum of (4.78) we could start the summation with $s = 0$. It is so, because the action of $Y_{k,0}^{[l+1]}$ on all components, starting from the second one, can be compensated by that of $Y_{k,s}^{[l+1]}$, for $s \geq 1$, and the action of $Y_{k,0}^{[l+1]}$ on the first component is irrelevant since the first component can be arbitrary in any feedforward system.

4.12.5 Feedforward and Strict Feedforward Systems on $\mathbb{R}^4$

The aim of this subsection is to illustrate our results by comparing feedforward and strict feedforward systems in $\mathbb{R}^4$. Notice that this is the lowest dimension in which both classes are nontrivial since in $\mathbb{R}^3$ all systems with controllable linearization can be brought to the feedforward form via feedback.

4.12.5.1 Feedforward Case

Consider a system on $\mathbb{R}^4$ and assume that it is feedback equivalent to the feedforward form up to terms of degree m . Then, according to Theorem 27, we can assume, without loss of generality, that the system takes the form

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 + \sum_{k=2}^m \hat{f}_1^{[k]}(\xi) + \bar{f}_1^{[m+1]}(\xi) \\ \dot{\xi}_2 &= \xi_3 + \sum_{k=2}^m \hat{f}_2^{[k]}(\xi) + \bar{f}_2^{[m+1]}(\xi) \\ \dot{\xi}_3 &= \xi_4 \\ \dot{\xi}_4 &= u\end{aligned}\tag{4.79}$$

where the homogeneous vector fields $\hat{f}_1^{[k]}(\partial/\partial\xi_1) + \hat{f}_2^{[k]}(\partial/\partial\xi_2)$, for $2 \leq k \leq m$, are in feedforward normal form (see Definition 4) and

$$\begin{aligned}\bar{f}^{[m+1]} &= \bar{f}_1^{[m+1]} \frac{\partial}{\partial\xi_1} + \bar{f}_2^{[m+1]} \frac{\partial}{\partial\xi_2} \\ &= \left(\xi_3^2 P_{1,3}^{[m-1]}(\xi_1, \xi_2, \xi_3) + \xi_4^2 P_{1,4}^{[m-1]}(\xi_1, \xi_2, \xi_3, \xi_4) \right) \frac{\partial}{\partial\xi_1} \\ &\quad + \left(\xi_4^2 P_{2,4}^{[m-1]}(\xi_1, \xi_2, \xi_3, \xi_4) \right) \frac{\partial}{\partial\xi_2}\end{aligned}$$

is in the Kang normal form $\Sigma^{[m+1]}$. Let us consider, for any $1 \leq s \leq m$, the following homogeneous vector field

$$Y_{2,s}^{[m]} = \xi_1^s \xi_2^{m-s} \frac{\partial}{\partial\xi_2} + L_{A\xi} (\xi_1^s \xi_2^{m-s}) \frac{\partial}{\partial\xi_3} + L_{A\xi}^2 (\xi_1^s \xi_2^{m-s}) \frac{\partial}{\partial\xi_4}$$

and construct the corresponding homogeneous system

$$\Sigma_{2,s}^{[m+1]} : \dot{\xi} = A\xi + Bu + \left[\bar{f}^{[2]}, Y_{2,s}^{[m]} \right](\xi)$$

where the Kang–Krener quadratic normal form $\bar{f}^{[2]}$ is given by

$$\bar{f}^{[2]} = (a\xi_3^2 + b\xi_4^2) \frac{\partial}{\partial \xi_1} + c\xi_4^2 \frac{\partial}{\partial \xi_2}$$

Let us denote by $a_{2,s}^{[m+1]j,i+2}$ and $a^{[m+1]j,i+2}$, for $(j,i) \in \Delta$, the homogeneous invariants associated, respectively, to the homogeneous system $\Sigma_{2,s}^{[m+1]}$ and to the normal form

$$\Sigma_{\text{NF}}^{[m+1]} : \dot{\xi} = A\xi + Bu + \bar{f}^{[m+1]}(\xi)$$

Here (A, B) is the Brunovsky canonical form of dimension 4. Denote $C_2 = (0, 1, 0, 0)$. We have $\Delta = \{(1, 0); (1, 1); (2, 0)\}$. Since the only term that is not feedforward is present if $P_{2,4}^{[m-1]}$ depends on ξ_1 , we will focus our attention only on the invariants $a_{2,s}^{[m+1]2,2}$ and $a^{[m+1]2,2}$ given, respectively, by

$$a_{2,s}^{[m+1]2,2} = C_2 \text{ad}_B^2 [\bar{f}^{[2]}, Y_{2,s}^{[m]}] \quad \text{and} \quad a^{[m+1]2,2} = C_2 \text{ad}_B^2 \bar{f}^{[m+1]} = \frac{\partial^2 \xi_4^2 P_{2,4}^{[m-1]}}{\partial \xi_4^2}$$

A direct computation gives

$$a_{2,s}^{[m+1]2,2} = 2bs\xi_1^{s-1}\xi_2^{m-s} - 2c(m-s)\xi_1^s\xi_2^{m-s-1}$$

Then the system (4.79) is feedback equivalent to the feedforward form up to order $m+1$ if and only if there exist real constants $\sigma_{2,s}$, for $1 \leq s \leq m-1$, such that

$$a^{[m+1]2,2} - \sum_{s=1}^m \sigma_{2,s} a_{2,s}^{[m+1]2,2} = Q^{[m-1]}(\xi_2, \xi_3, \xi_4)$$

for some homogeneous polynomial $Q^{[m-1]}(\xi_2, \xi_3, \xi_4)$, which is equivalent to the condition that

$$P_{2,4}^{[m-1]}(\xi) = S_1^{[m-1]}(\xi_1, \xi_2) + S_2^{[m-1]}(\xi_2, \xi_3, \xi_4)$$

The codimension $c_{\text{FF}}(m+1)$ of the space of homogeneous systems of degree $m+1$, which are feedback equivalent to feedforward form, is equal to the dimension of the space of all homogeneous polynomials of degree $m-1$ of the form

$$Q_2^{[m-1]}(\xi) = \xi_1 \xi_3 R_1^{[m-3]}(\xi_1, \xi_2, \xi_3) + \xi_1 \xi_4 R_2^{[m-3]}(\xi_1, \xi_2, \xi_3, \xi_4)$$

We thus get

$$c_{\text{FF}}(m+1) = \frac{(m-1)(m-2)}{2} + \frac{(m-2)(m-1)m}{6} = \frac{(m+3)(m-1)(m-2)}{6}$$

4.12.5.2 Strict Feedforward Case

Now, consider a system on $\mathbb{R}^4$ and assume that it is feedback equivalent to a strict feedforward form up to terms of degree m . Then, according to Theorem 27, we can assume, without loss of generality, that the system takes the form (4.79), where the homogeneous vector fields $\hat{f}_1^{[k]}(\partial/\partial\xi_1) + \hat{f}_2^{[k]}(\partial/\partial\xi_2)$, for $2 \leq k \leq m$, are in strict feedforward normal form (see Definition 3) and $\bar{f}^{[m+1]}$ in Kang normal form. We assume $m_0 = 2$ which yields $l = m - 1$, and we consider the vector field

$$\begin{aligned} Y_{2,0}^{[m]} &= \xi_2^m \frac{\partial}{\partial\xi_2} + L_{A\xi}(\xi_2^m) \frac{\partial}{\partial\xi_3} + L_{A\xi}^2(\xi_2^m) \frac{\partial}{\partial\xi_4} \\ &= \xi_2^m \frac{\partial}{\partial\xi_2} + m\xi_2^{m-1}\xi_3 \frac{\partial}{\partial\xi_3} + \left(m(m-1)\xi_2^{m-2}\xi_3^2 + m\xi_2^{m-1}\xi_4\right) \frac{\partial}{\partial\xi_4} \end{aligned}$$

whose corresponding homogeneous system is

$$\Sigma_{2,0}^{[m+1]} : \dot{\xi} = A\xi + Bu + \left[\bar{f}^{[2]}, Y_{2,0}^{[m]}\right]$$

Denote by $a_{2,0}^{[m+1]j,i+2}$ and $\bar{a}^{[m+1]j,i+2}$, for $(j,i) \in \Delta = \{(1,0); (1,1); (2,0)\}$, the homogeneous invariants associated, respectively, to the homogeneous systems $\Sigma_{2,0}^{[m+1]}$ and to the normal form

$$\Sigma_{\text{NF}}^{[m+1]} : \dot{\xi} = A\xi + Bu + \bar{f}^{[m+1]}(\xi)$$

Denote $C_1 = (1, 0, 0, 0)$ and $C_2 = (0, 1, 0, 0)$. We will calculate the invariants $a_{2,0}^{[m+1]1,2}$, $a_{2,0}^{[m+1]2,2}$, and $a_{2,0}^{[m+1]1,3}$ of system $\Sigma_{2,0}^{[m+1]}$. We have

$$a_{2,0}^{[m+1]1,2} = C_1 \text{ad}_B^2 [\bar{f}^{[2]}, Y_{2,0}^{[m]}], \quad a_{2,0}^{[m+1]2,2} = C_2 \text{ad}_B^2 [\bar{f}^{[2]}, Y_{2,0}^{[m]}]$$

and

$$a_{2,0}^{[m+1]1,3} = C_1 \left(\text{ad}_{AB} X_2^{[m]} - \text{ad}_{A^2B} X_1^{[m]} \right) (\pi_3(\xi))$$

where $\pi_3(\xi) = (\xi_1, \xi_2, \xi_3)^T$, $X_1^{[m]} = \text{ad}_B[\bar{f}^{[2]}, Y_{2,0}^{[m]}]$, and $X_2^{[m]} = \text{ad}_{AB}[\bar{f}^{[2]}, Y_{2,0}^{[m]}] - \text{ad}_{A\xi} \text{ad}_B[\bar{f}^{[2]}, Y_{2,0}^{[m]}]$.

A direct computation gives

$$a_{2,0}^{[m+1]1,2} = a_{2,0}^{[m+1]2,2} = -2cm\xi_2^{m-1}$$

and

$$a_{2,0}^{[m+1]1,3} = -4am\xi_2^{m-1} + 8bm(m-1)(m-2)\xi_2^{m-3}\xi_3^2 - 4cm(m-1)\xi_2^{m-2}\xi_3$$

On the other hand, the invariants associated to the normal form are

$$\bar{a}^{[m+1]1,2} = \text{ad}_B^2 \bar{f}_1^{[m+1]} = \frac{\partial^2 \xi_4^2 P_{1,4}^{[m-1]}}{\partial \xi_4^2}$$

$$\bar{a}^{[m+1]1,3} = \text{ad}_{AB}^2 \bar{f}_1^{[m+1]}(\pi_3(\xi)) = \frac{\partial^2 \xi_3^2 P_{1,3}^{[m-1]}}{\partial \xi_3^2}(\pi_3(\xi))$$

and

$$\bar{a}^{[m+1]2,2} = \text{ad}_B^2 \bar{f}_2^{[m+1]} = \frac{\partial^2 \xi_4^2 P_{2,4}^{[m-1]}}{\partial \xi_4^2}$$

Then the system (4.79) is feedback equivalent to a strict feedforward form up to order $m + 1$ if and only if there exist a real constants $\sigma_{2,0}$ such that

$$\bar{a}^{[m+1]1,2} - \sigma_{2,0} a_{2,0}^{[m+1]1,2} = Q_{1,4}^{[m-1]}(\xi_2, \xi_3, \xi_4)$$

$$\bar{a}^{[m+1]1,3} - \sigma_{2,0} a_{2,0}^{[m+1]1,3} = Q_{1,3}^{[m-1]}(\xi_2, \xi_3)$$

and

$$\bar{a}^{[m+1]2,2} - \sigma_{2,0} a_{2,0}^{[m+1]2,2} = Q_{2,4}^{[m-1]}(\xi_3, \xi_4)$$

for some homogeneous polynomials $Q_{1,4}^{[m-1]}(\xi_2, \xi_3, \xi_4)$, $Q_{1,3}^{[m-1]}(\xi_2, \xi_3)$, and $Q_{2,4}^{[m-1]}(\xi_3, \xi_4)$. These conditions are equivalent to the fact that

$$P_{1,4}^{[m-1]}(\xi) = S_{1,4}^{[m-1]}(\xi_2, \xi_3, \xi_4), \quad P_{1,3}^{[m-1]}(\xi) = S_{1,3}^{[m-1]}(\xi_2, \xi_3), \quad \text{and}$$

$$P_{2,4}^{[m-1]}(\xi) = \lambda \xi_2^{m-1} + S_{2,4}^{[m-1]}(\xi_3, \xi_4)$$

The codimension $c_{\text{SFF}}(m + 1)$ of the space of homogeneous systems of degree $m + 1$, which are feedback equivalent to strict feedforward form, is equal to the dimension of the space of all homogeneous vector fields of degree $m - 1$ of the form

$$\left(\xi_1 \xi_3^2 R_{1,3}^{[m-3]}(\xi_1, \xi_2, \xi_3) + \xi_1 \xi_4^2 R_{1,4}^{[m-2]}(\xi_1, \xi_2, \xi_3, \xi_4) \right) \frac{\partial}{\partial \xi_1}$$

$$+ \xi_4^2 \left(\xi_1 R_{2,4}^{[m-2]}(\xi_1, \xi_2, \xi_3, \xi_4) + \xi_2 \bar{R}_{2,4}^{[m-2]}(\xi_2, \xi_3, \xi_4) \right) \frac{\partial}{\partial \xi_2}$$

with $\bar{R}_{2,4}^{[m-2]}(\xi_2, 0, 0) = 0$, that is, $\bar{R}_{2,4}^{[m-2]}(\xi_2, 0, 0) = \xi_3 \tilde{R}_{2,4}^{[m-3]}(\xi_2, \xi_3) + \xi_4 \hat{R}_{2,4}^{[m-3]}(\xi_2, \xi_3, \xi_4)$.

We thus get

$$\begin{aligned} c_{\text{SFF}}(m+1) &= \frac{(m+1)m(m-1)}{6} + \frac{m(m-1)}{2} + \frac{(m+1)m(m-1)}{6} \\ &\quad + \frac{m-2}{1} + \frac{(m-1)(m-2)}{2} \\ &= \frac{m^3 + 3m^2 - 4m - 3}{3} \end{aligned}$$

REMARK 4

The codimensions $c_{\text{FF}}(m+1)$ and $c_{\text{SFF}}(m+1)$ are computed for *generic systems*, in particular in the case when $\bar{f}_2^{[2]} \neq 0$. If $\bar{f}_2^{[2]} = 0$, then those codimensions are modified as follows

$$\begin{aligned} \tilde{c}_{\text{FF}}(m+1) &= c_{\text{FF}}(m+1) + 1 = \frac{(m+3)(m-1)(m-2)}{6} + 1 \\ \tilde{c}_{\text{SFF}}(m+1) &= c_{\text{SFF}}(m+1) + 1 = \frac{m^3 + 3m^2 - 4m}{3} \end{aligned}$$

In each case, the gap between the two codimensions is equal to $(m^3 + 6m^2 - m - 12)/6$.

Finally, we will compute the codimension of the space of linearizable homogenous systems of degree $m+1$ on $\mathbb{R}^4$. The homogeneous part of degree $m+1$ of (4.79) is given by two polynomials $P_{1,4}$ and $P_{2,4}$ of four variables and one polynomial $P_{1,3}$ of three variables. Linearizability is equivalent to vanishing of all of them and thus the codimension of linearizable homogeneous systems is

$$c_{\text{FL}}(m+1) = \frac{(m-1)m}{2} + \frac{2(m-1)m(m+1)}{6} = \frac{(2m+5)(m-1)m}{6}$$

4.12.6 Geometric Characterization of Feedforward and Strict Feedforward Systems

In the previous subsection, we proposed a step-by-step constructive method to bring a system into a feedforward form and strict feedforward form whenever possible [80, 82, 84]. The problem of transforming a system, affine with respect to controls, into (strict) feedforward form via a diffeomorphism (i.e., via a nonlinear change of coordinates), was studied earlier [63]. A geometric description of systems in feedforward form has been given elsewhere [4]. These conditions, although being intrinsic, are not checkable.

In the present section, we look at the problem in the spirit of Ref. [4] but we focus our attention on vector fields rather than on invariant

distributions. It turns out that feedback equivalence (resp. state-space equivalence) to the strict feedforward form can be characterized by the existence of a sequence of infinitesimal symmetries (resp. strong infinitesimal symmetries) of the system.

4.12.7 Symmetries and Strict Feedforward Form

In this section, we will establish results relating symmetries and strict feedforward forms. To start with, recall (see Introduction) two basic notions of equivalence of control systems. The word smooth will mean throughout C^∞ -smooth and all control systems are assumed to be smooth.

Two control systems

$$\Sigma : \dot{x} = f(x) + g(x)u, \quad x \in X$$

and

$$\tilde{\Sigma} : \dot{\tilde{x}} = \tilde{f}(\tilde{x}) + \tilde{g}(\tilde{x})\tilde{u}, \quad \tilde{x} \in \tilde{X}$$

are called S-equivalent, if there exists a smooth diffeomorphism $\phi : X \rightarrow \tilde{X}$, such that

$$\phi_* f = \tilde{f} \quad \text{and} \quad \phi_* g = \tilde{g}$$

(we take $u = \tilde{u}$), and they are called F-equivalent (feedback equivalent), if there exists a smooth diffeomorphism $\phi : X \rightarrow \tilde{X}$ and smooth functions α, β , satisfying $\beta(\cdot) \neq 0$, such that

$$\phi_*(f + g\alpha) = \tilde{f} \quad \text{and} \quad \phi_*(g\beta) = \tilde{g}$$

Recall also (see Section 4.11.4) that a vector field v on an open subset $X \subset \mathbb{R}^n$ is an *infinitesimal symmetry* of the system Σ if the (local) flow γ_t^v of v is a local symmetry of Σ , for any t for which it exists.

We will also be dealing with the following stronger notions. A diffeomorphism $\phi : X \rightarrow \tilde{X}$ is a *strong symmetry* of Σ if it preserves the vector fields f and g (and not only the affine distribution $\mathcal{A}$ spanned by them), that is, if

$$\phi_* f = f \quad \text{and} \quad \phi_* g = g$$

A local strong symmetry is a local diffeomorphism preserving f and g . We say that a vector field v on an open subset $X \subset \mathbb{R}^n$ is an *infinitesimal strong symmetry* of the system Σ if the (local) flow γ_t^v of v is a local strong symmetry of Σ , for any t for which it exists.

Consider the system Σ and denote by $\mathcal{G}$ the distribution spanned by the vector field g . We have the following characterization of infinitesimal symmetries and strong symmetries.

PROPOSITION 16

1. A vector field v is an infinitesimal strong symmetry of Σ if and only if

$$[v, g] = 0, \quad \text{and} \quad [v, f] = 0$$

2. A vector field v , such that $v(p) \neq 0$, is an infinitesimal symmetry of Σ , locally at p , if and only if

$$[v, g] = 0 \bmod \mathcal{G}, \quad \text{and} \quad [v, f] = 0 \bmod \mathcal{G}$$

in a neighborhood of p .

The second item remains true even if $g(p) = 0$. In this case, we have to understand $\mathcal{G}$ as being the module of vector fields generated by g over the ring of smooth functions. An infinitesimal symmetry v is called stationary at $p \in X$ if $v(p) = 0$ and nonstationary if $v(p) \neq 0$.

Assume that v is a strong infinitesimal symmetry of Σ , nonstationary at $p \in X$. Then there exist a neighborhood X_p of p and the factor system $\Sigma/\sim_v$, where the equivalence relation $\sim_v$ is induced by the local action of the 1-parameter local group defined by v , that is, $q_1 \sim_v q_2$ if and only if they belong to the same integral curve of v (more precisely, to the same connected component of the intersection of an integral curve of v with X_p).

THEOREM 31

The following condition are equivalent.

1. Σ is, locally at $p \in X$, S -equivalent to the affine strict feedforward form.
2. Each system $\Sigma_1, \Sigma_2, \dots, \Sigma_n$ possesses a strong infinitesimal nonstationary symmetry v_i , where Σ_1 is the restriction of Σ to a neighborhood X_p and

$$\Sigma_{i+1} = \Sigma_i / \sim_{v_i}$$

with $\sim_{v_i}$ the equivalence relation defined by the local action of the 1-parameter group of v_i .

3. There exist smooth vector fields $w_1, \dots, w_n$, independent at $p \in X$, such that, locally at p ,

$$[w_i, w_j] \in \mathcal{D}_{i-1}, \quad [w_i, g] \in \mathcal{D}_{i-1}, \quad [w_i, f] \in \mathcal{D}_{i-1}$$

for any $1 \leq i \leq n$ and $j \leq i$, where $\mathcal{D}_0 = 0$ and $\mathcal{D}_i = \text{span} \{w_1, \dots, w_i\}$.

4. There exist smooth vector fields $\tilde{w}_1, \dots, \tilde{w}_n$, independent at $p \in X$, such that, locally at p ,

$$[\tilde{w}_i, \tilde{w}_j] = 0, \quad [\tilde{w}_i, g] \in \tilde{\mathcal{D}}_{i-1}, \quad [\tilde{w}_i, f] \in \tilde{\mathcal{D}}_{i-1}$$

for any $1 \leq i \leq n$ and $j \leq i$, where $\tilde{\mathcal{D}}_0 = 0$ and $\tilde{\mathcal{D}}_i = \text{span} \{\tilde{w}_1, \dots, \tilde{w}_i\}$.

In Section 4.12.8, we will show that the problem of transforming a general system to SFF can be reduced to Theorem 31 by a preintegration. A detailed proof of Theorem 31 is given elsewhere [73].

Theorem 31 implies that an invariant characterization of the affine strict feedforward form involves vector fields (forming a sequence of infinitesimal symmetries) rather than invariant distributions. To be more precise, a characterization of the *affine feedforward form*

$$\begin{aligned}
 \dot{z}_1 &= f_1(z_1, \dots, z_n) + g_1(z_1, \dots, z_n)u \\
 \dot{z}_2 &= f_2(z_2, \dots, z_n) + g_2(z_2, \dots, z_n)u \\
 &\vdots \\
 \dot{z}_{n-1} &= f_{n-1}(z_{n-1}, z_n) + g_{n-1}(z_{n-1}, z_n)u \\
 \dot{z}_n &= f_n(z_n) + g_n(z_n)u
 \end{aligned} \tag{AFF}$$

was obtained by Astolfi and Mazenc [4] in terms of invariant distributions as given in the following proposition.

PROPOSITION 17

The system Σ is locally equivalent to the affine feedforward form if and only if there exists a sequence of distributions

$$\mathcal{D}_1 \subset \dots \subset \mathcal{D}_n$$

where $\mathcal{D}_i$ is involutive and of rank i , such that

$$[\mathcal{D}_i, g] \subset \mathcal{D}_i, \quad [\mathcal{D}_i, f] \subset \mathcal{D}_i$$

A first guess for a characterization of the affine strict feedforward form could be [4] the existence of a nested sequence of involutive distributions $\mathcal{D}_i$, of constant rank i , satisfying

$$[\mathcal{D}_i, g] \subset \mathcal{D}_{i-1}, \quad [\mathcal{D}_i, f] \subset \mathcal{D}_{i-1}$$

This is not a correct answer for two reasons. First, the latter conditions are not invariant, that is, even if they are satisfied for some vector fields $w_1, \dots, w_i$ spanning $\mathcal{D}_i$ then, in general, for other generators of the same distribution $\mathcal{D}_i$, we will have on the right the inclusion in $\mathcal{D}_i$ (and not in $\mathcal{D}_{i-1}$). Secondly, the aforementioned conditions, even reformulated in terms of vector fields, are not sufficient for equivalence to affine strict feedforward

form. Indeed, the condition that there exist linearly independent vector fields $w_1, \dots, w_n$ such that

$$[w_i, g] \in \mathcal{D}_{i-1}, \quad [w_i, f] \in \mathcal{D}_{i-1}$$

for any $1 \leq i \leq n$, where $\mathcal{D}_0 = 0$ and $\mathcal{D}_i = \text{span} \{w_1, \dots, w_i\}$ are involutive, does not imply S -equivalence to the affine strict feedforward form unless we assume an additional property on the w_i 's: like the first condition of Theorem 31 (3) (which is the weakest possible) or the first condition of Theorem 31 (4), which is the strongest one.

We have an analogous result for feedback equivalence to the strict feedforward form, where the role of strong infinitesimal symmetries is replaced by that of infinitesimal symmetries. To state this, we need the following considerations. We will write $\Sigma(f, g)$, to denote the system Σ defined by the pair of vector fields (f, g) . Assume that v is an infinitesimal symmetry of $\Sigma(f, g)$, nonstationary at $p \in X$, that is, such that $v(p) \neq 0$. Then the second item of Proposition 16 implies that there exists a feedback pair (α, β) such that v is a strong infinitesimal symmetry of the system $\tilde{\Sigma}(\tilde{f}, \tilde{g})$, where $\tilde{f} = f + g\alpha$ and $\tilde{g} = g\beta$. Thus there exists a neighborhood X_p of p in which the factor system $\tilde{\Sigma}/\sim_v$ is well defined, where the equivalence relation $\sim_v$ is induced by the local action of the 1-parameter local group defined by v . Notice that given a system Σ , there are many systems $\tilde{\Sigma}(\tilde{f}, \tilde{g})$, feedback equivalent to Σ , and such that v is a strong infinitesimal symmetry of $\tilde{\Sigma}$. We will denote by $\tilde{\Sigma}$ any of those systems. Actually, any two such systems are equivalent by a feedback pair $(\tilde{\alpha}, \tilde{\beta})$, where the functions $\tilde{\alpha}$ and $\tilde{\beta}$ are constant on the trajectories of v .

THEOREM 32

The following condition are equivalent:

1. Σ is, locally at $p \in X$, F -equivalent to the affine strict feedforward form satisfying $g_n \neq 0$.
2. Each system $\Sigma_1, \Sigma_2, \dots, \Sigma_n$ possesses an infinitesimal symmetry v_i , where Σ_1 is the restriction of Σ to a neighborhood X_p and

$$\Sigma_{i+1} = \tilde{\Sigma}_i / \sim_{v_i}$$

where $\sim_{v_i}$ is the equivalence relation induced by the local action of the 1-parameter group of v_i , and such that v_i and the control vector field g_i of Σ_i are independent, for $1 \leq i \leq n-1$.

3. There exist smooth vector fields $w_1, \dots, w_n$, independent at $p \in X$, such that, locally at p ,

$$[w_i, w_j] \in \mathcal{D}_{i-1}, \quad [w_i, g] \in \mathcal{D}_{i-1} + \mathcal{G}, \quad [w_i, f] \in \mathcal{D}_{i-1} + \mathcal{G}$$

for any $1 \leq i \leq n$ and $j \leq i$, where $\mathcal{D}_0 = 0$ and $\mathcal{D}_i = \text{span} \{w_1, \dots, w_i\}$ and, moreover, $g(p) \notin \mathcal{D}_{n-1}(p)$.

4. There exist smooth vector fields $\tilde{w}_1, \dots, \tilde{w}_n$, independent at $p \in X$, such that, locally at p ,

$$[\tilde{w}_i, \tilde{w}_j] = 0, \quad [\tilde{w}_i, g] \in \tilde{\mathcal{D}}_{i-1} + \mathcal{G}, \quad [\tilde{w}_i, f] \in \tilde{\mathcal{D}}_{i-1} + \mathcal{G}$$

for any $1 \leq i \leq n$ and $j \leq i$, where $\tilde{\mathcal{D}}_0 = 0$ and $\tilde{\mathcal{D}}_i = \text{span} \{\tilde{w}_1, \dots, \tilde{w}_i\}$ and, moreover, $g(p) \notin \tilde{\mathcal{D}}_{n-1}(p)$.

The assumption $g(p) \notin \mathcal{D}_{n-1}(p)$ can be dropped (equivalently, we allow for $g_n = 0$) if we understand the conditions (3) and (4) as well as those of the second item of Proposition 16 in the sense of module of vector fields and not of distributions.

A proof of Theorem 32 follows the same line as that of Theorem 31, the only difference is to show that in the successive steps, the existence of infinitesimal symmetries does not depend on the choice of $\tilde{\Sigma}_i$ in $\Sigma_{i+1} = \tilde{\Sigma}_i / \sim_{v_i}$.

4.12.8 Strict Feedforward Form: Affine Versus General

In this section, we will show that the problem of transforming a general control system to the strict feedforward form can be reduced to that for affine systems by taking the preintegration. The same procedure of extension (compare Proposition 1) has been already used for the problems of linearization and decoupling [94] and equivalence to the p -normal form [71].

Consider a general nonlinear control system

$$\Pi : \dot{x} = f(x, u)$$

where $x \in X$, an open subset of $\mathbb{R}^n$, $u \in \mathbb{R}$. Together with Σ , we consider its extension (preintegration)

$$\Pi^e : \dot{x}^e = f^e(x^e) + g^e(x^e)w$$

where $x^e = (x, u) \in X \times \mathbb{R}^1$, $w \in \mathbb{R}$, and the dynamics are given by $f^e(x^e) = (f(x, u), 0)^T$ and $g^e(x^e) = (\partial/\partial u)$. Notice that Σ^e is a control-affine system controlled by the derivative $\dot{u} = w$ of the original control u .

Recall that $\mathcal{L}_0$ denotes the Lie ideal generated by $\{f_u - f_{\bar{u}}\}$, $u, \bar{u} \in U$, in the Lie algebra $\mathcal{L}$ of the system Π . Assume that $\dim \mathcal{L}_0(p) = n$.

PROPOSITION 18

The system Π is S -equivalent (resp. F -equivalent), locally at (x_0, u_0) , to the strict feedforward form if and only if the extension Π^e is, locally at

$x_0^e = (x_0, u_0)$, S -equivalent (resp. F -equivalent) to the affine strict feedforward form.

The proof is based on showing that a diffeomorphism bringing Π^e into the affine strict feedforward form is of a special form: states depend on states only and the control is preserved. In particular, we show the following statement, which is of independent interest.

COROLLARY 11

If the system Σ is in an affine strict feedforward form satisfying $g_n \neq 0$, then it is S -equivalent to another affine strict feedforward form, for which $g_1 = \dots = g_{n-1} = 0$.

4.12.9 Strict Feedforward Systems on the Plane

In this section, we will describe strict feedforward systems on the plane. Consider a system Σ on an open subset X of $\mathbb{R}^2$ and suppose that $g(p) \neq 0$. We define the *multiplicity* of Σ at p as the smallest integer μ , such that g and $\text{ad}_g^\mu f$ are linearly independent at p . Notice that the notion of multiplicity is feedback invariant [46]. If the multiplicity is $\mu = 1$, then the system is feedback linearizable and thus feedback equivalent to the affine strict feedforward form. The case of multiplicity $\mu \geq 2$ is described by the following proposition.

PROPOSITION 19

Consider a system Σ on open subset X of $\mathbb{R}^2$ and suppose that $g(p) \neq 0$ and that it has multiplicity $\mu \geq 2$ at p .

1. *If f and g are linearly dependent at p , then Σ is locally F -equivalent to the strict feedforward form if and only*

$$f = \gamma \text{ad}_g f \bmod \mathcal{G}$$

where γ is a smooth function such that the smooth function φ defined by

$$f = \varphi \text{ad}_g^\mu f \bmod \mathcal{G}$$

is divisible by γ^μ . Moreover, in this case Σ is locally F -equivalent to

$$\begin{aligned}\dot{z}_1 &= z_2^\mu \\ \dot{z}_2 &= v\end{aligned}$$

2. If f and g are linearly independent at p , then Σ is locally F -equivalent to the strict feedforward form if and only

$$\text{ad}_g f = \gamma \text{ad}_g^2 f \bmod \mathcal{G}$$

where γ is a smooth function such that the smooth function ψ defined by

$$\text{ad}_g f = \psi \text{ad}_g^\mu f \bmod \mathcal{G}$$

is divisible by $\gamma^{\mu-1}$. Moreover in this case Σ is locally F -equivalent to

$$\dot{z}_1 = 1 + \epsilon z_2^\mu$$

$$\dot{z}_2 = v$$

It has been proved earlier [46] that any planar system with a finite multiplicity μ at p is locally feedback equivalent to the following system around $0 \in \mathbb{R}^2$:

$$\begin{aligned} \dot{z}_1 &= z_2^\mu + a_{\mu-2} z_2^{\mu-2} + \cdots + a_1 z_2 + a_0 \\ \dot{z}_2 &= v \end{aligned}$$

where the smooth functions a_i , for $0 \leq i \leq \mu - 2$, depend on z_1 only and satisfy $a_i(0) = 0$ (except for a_0 in the case f and g independent at p). Moreover, we can always normalize one of the functions a_i (in particular, we can take $a_0 = \pm 1$ if $a_0(0) \neq 0$) and then the infinite jets of all remaining functions are feedback invariant. Proposition 19 implies that among all planar system only those are F -equivalent to the affine strict feedforward form for which all the above invariants are identically zero.

4.13 Analytic Normal Forms: A Class of Strict Feedforward Systems

In the previous sections, we have developed the theory of feedback classification following a formal approach introduced by Kang and Krener. Although the normal forms obtained are formal, the theory has proved to be very useful in analyzing structural properties of nonlinear control systems. It has been used to study bifurcations of nonlinear systems [52, 55], has led to a complete description of symmetries around equilibrium, presented in Section 4.11 [72, 73], and allowed to characterize systems equivalent to feedforward and strict feedforward forms (see Section 4.12) [80, 82, 84].

A natural question to ask is whether normal and canonical forms presented in earlier sections are convergent. It is already known that the problem of convergence is difficult even for dynamical systems whose convergence depend on the location of the eigenvalues of the linear part. Those eigenvalues stand to be invariants for dynamical systems which is a first difference with control systems because the notion of eigenvalues is meaningless. It has been proved (see [3, 12]) that a dynamical system is biholomorphically equivalent to its linear part if the spectrum of its linearization is not resonant and belongs either to the Poincaré domain or to the Siegel domain with type (C, ν) . When the spectrum is resonant and belongs to the Poincaré domain, then the Poincaré–Dulac theorem shows that the dynamical system is biholomorphically equivalent to a polynomial vector field. We will not explicitly recall those results here because of space limitations and we refer the reader to the existing literature.

For control systems, Kang [50] derived from Ref. [56], and Ref. [57] (see also [38] and Proposition 7) that if an analytic control system is linearizable by a formal transformation, then it is linearizable by an analytic transformation. Kang [50] also gives a class of nonlinearizable three-dimensional analytic control systems which are equivalent to their normal forms by analytic transformations. Those are the only results about convergence of the step-by-step normalizing transformations known to us to this date (see, however, the C^∞ -smooth and/or analytic normal forms of [9], [40], [46], [69], [75], [95]).

In this section, we study a class of nonlinear systems called special strict feedforward forms, and we show that this class could be brought to its normal form (actually canonical form) via analytic transformations.

Consider an *analytic* single-input control system $\dot{x} = f(x, u)$, in strict feedforward form, that is, such that $f_j(x, u) = f_j(x_{j+1}, \dots, x_n, u)$, $1 \leq j \leq n$.

Notice that each component decomposes uniquely as

$$f_j(x, u) = a_j(x_{j+1}) + F_j(x_{j+1}, \dots, x_n, u), \quad \text{with } F_j(x_{j+1}, 0, \dots, 0) = 0 \quad (4.80)$$

A (*special strict feedforward form*) SSFF is an analytic strict feedforward form for which

$$a_j(x_{j+1}) = k_j x_{j+1}, \quad \text{whenever } \frac{\partial a_j}{\partial x_{j+1}}(0) \neq 0 \quad (4.81)$$

The main result of this section is given in the following theorem [87].

THEOREM 33

Consider an analytic special strict feedforward form given by (4.80) and (4.81). There exists an analytic feedback transformation that brings the system (4.80) and

(4.81) into the normal form

$$\Pi_{\text{SSFN}} : \begin{cases} \dot{z}_1 = z_2 + \sum_{i=3}^{n+1} z_i^2 P_{1,i}(z_2, \dots, z_i) \\ \vdots \\ \dot{z}_j = z_{j+1} + \sum_{i=j+2}^{n+1} z_i^2 P_{j,i}(z_{j+1}, \dots, z_i) \\ \vdots \\ \dot{z}_{n-1} = z_n + z_{n+1}^2 P_{n-1,n+1}(z_{n+1}) \\ \dot{z}_n = v \end{cases} \quad (4.82)$$

where $P_{j,i}(z_{j+1}, \dots, z_i)$ are analytic functions of the indicated variables, and $z_{n+1} = v$.

The main remark is that the normal form itself is in a strict feedforward form. Moreover, this normal form coincides with the canonical form defined elsewhere [83]. This leads to the following theorem.

THEOREM 34

Two special strict feedforward systems $(\text{SSFF})_1$ and $(\text{SSFF})_2$ are feedback equivalent if and only if their normal forms Π_{SSFN} are equal (after possible reparametrization $z_i = \lambda x_i$).

The proof of Theorem 33 is detailed elsewhere [87]. The proof of Theorem 34 follows automatically after normalization of the first nonlinearizable homogeneous vector field because no components of the special strict feedforward forms depend on the first variable x_1 .

It has been proved earlier [87] that special strict feedforward forms define the only class of strict feedforward systems that can be brought to a normal form, still being in strict feedforward form.

Indeed, if

$$\dot{z} = \tilde{f}(z, v)$$

is another analytic strict feedforward form, that is, such that

$$\tilde{f}_j(z, u) = \tilde{a}_j(z_{j+1}) + \tilde{F}_j(z_{j+1}, \dots, z_n, v), \quad \text{with } \tilde{F}_j(z_{j+1}, 0, \dots, 0) = 0 \quad (4.83)$$

for any $1 \leq j \leq n$, then we have

THEOREM 35

The system (4.83) is feedback equivalent to a special strict feedforward form if and only if

$$\tilde{a}_j(z_{j+1}) = \tilde{k}_j z_{j+1}, \quad \text{whenever } \frac{\partial \tilde{a}_j}{\partial z_{j+1}}(0) \neq 0$$

that is, the system is in special strict feedforward form in its coordinates.

Whether it is possible to bring any strict feedforward system into its normal form Π_{NF} by analytic transformation is unclear yet, but if true, the normal form will no longer be strict feedforward. We refer the reader elsewhere [87] for more details.

4.14 Conclusions

This chapter, as its name indicates, is an attempt to summarize the diverse results about normal forms obtained in the past two decades using a formal approach. Starting from the pioneer work of Poincaré (Section 4.2), we then described, in Section 4.3, normal forms for single-input control systems obtained by Kang and Krener using classical Poincaré's technique. The work of Kang and Krener was completed by the authors to obtain canonical forms, dual normal, and dual canonical forms. Those results were discussed, respectively, in Section 4.4 to Section 4.6. Results in previously mentioned sections concern single-input control systems with controllable linearization. The uncontrollable linearization case as well as the multi-input case came as a generalization of previous results, respectively, in Section 4.7 and Section 4.8. In each of those sections, the results obtained have been compared to results in earliest sections and have been shown to be their generalizations. The feedback linearization of control systems was originally the first problem dealt with in feedback classification. However, we have chosen to introduce the results regarding this only in Section 4.9 in order to make a parallelism between the formal approach that provides a step-by-step procedure of linearization and the classical approach using distributions. Although the main results concern continuous time control systems, it would be unfair not to mention the discrete time case. Thus we have devoted Section 4.10 to normal forms of discrete time control systems. This formal approach, introduced by Kang and Krener for control systems, has proved to be very useful in analyzing systems. If we had enough space, we could widen our chapter, among other topics, to bifurcations, stabilization, and observability of control systems. A complete description of symmetries of control systems has been obtained by the authors using this formal approach, and these results are described in Section 4.11. The same approach has led to a step-by-step characterization of systems feedback equivalent to feedforward systems or to strict feedforward systems. For each degree of homogeneity, necessary and sufficient conditions were obtained and presented in Section 4.12. The amazing point about these notions is that they are all nicely related and have the same roots: normal and canonical forms. Indeed, we have shown that symmetries are described by canonical forms either in the analytic or formal category.

On the other hand, feedforward and strict feedforward systems are geometrically characterized in Section 4.12 using symmetries. Finally, in Section 4.13, it turns out that one of the biggest class (ever found) of control systems that could be brought to a normal and canonical form, using analytic transformations, is a subclass of strict feedforward systems. The important number of references listed in this chapter illustrates the interest in the notions presented here, and we may have certainly omitted many others. The interested reader will probably complete our work in this matter.

References

1. A.A. Agrachev and R.V. Gamkrlidze, Feedback invariant optimal control theory and differential geometry, I. Regular extremals, *J. Dynamical and Control Systems*, (3) 1997, pp. 343–389.
2. A.A. Agrachev, Feedback invariant optimal control theory and differential geometry, II. Jacobi curves for singular extremals, *J. Dynamical and Control Systems*, 4(1998), pp. 583–604.
3. V.I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, 2nd ed., Springer-Verlag, 1988.
4. A. Astolfi and F. Mazenc, A geometric characterization of feedforward forms, in *Proceedings of the MTNS'2000*, Perpignan, France, 2000.
5. S. Battilotti, Semiglobal stabilization of uncertain block-feedforward forms via measurement feedback, in *Proceedings of the NOLCOS'98*, Enschede, the Netherlands, 1998, pp. 342–347.
6. J.-P. Barbot, S. Monaco, and D. Normand-Cyrot Quadratic forms and approximative feedback linearization in Discrete Time, *Int. J. Control*, 67, 567–586, 1997.
7. V. Bogaevski and A. Povzner, *Algebraic Methods in Nonlinear Physics*, Springer-Verlag, New York, 1991.
8. B. Bonnard, Feedback equivalence for nonlinear systems and the time optimal control problem, *SIAM J. Control Optim.*, 29, 1300–1321, 1991.
9. B. Bonnard, Quadratic control systems, *Math. Control, Signals, Syst.*, 4, 1991, 139–160.
10. A. Bressan and F. Rampazzo, On differential systems with quadratic impulses and their applications to Lagrangian mechanics, *SIAM J. Control Optim.*, 31, 1205–1230, 1993.
11. R.W. Brockett, Feedback invariants for nonlinear systems, in *Proceedings of the IFAC Congress*, Helsinki, 1978.
12. A.D. Bruno, W. Hovingh, and C.S. Coleman, *The Local Method of Nonlinear Analysis of Differential Equations*, Springer Series in Soviet Mathematics, Springer, Berlin, 1989.
13. E. Cartan, *Théorie des groupes finis et continus et la géométrie différentielle traitées par la méthode du repère mobile*, Gauthier-Villars, Paris, 1937.

14. S. Čelikovsky and H. Nijmeijer, Equivalence of nonlinear systems to triangular form: the singular case, *Syst. Control Lett.*, 27, 135–144, 1996.
15. L. Cremean, W.B. Dunbar, D.V. Goh, J. Hickey, E. Kavins, J. Meltzer, and R. Murray, The Caltech Multi-Vehicle Wireless Testbed, *Proceedings of 40th CDC*, Las Vegas, USA, 2002.
16. E. Delaleau and W. Respondek, Lowering the orders of derivatives of controls in generalized state-space systems, *J. Math. Syst. Est. Contr.*, 8, 427–453, 1998.
17. W.B. Dunbar and R. Murray, Model predictive control of coordinated multi-vehicle formations, *Proceedings of 40th CDC*, Las Vegas, USA, 2002.
18. M. Flies, J. Lévine, and P. Rouchon, Generalized state variable representation for a simplified crane description, *Int. J. Control*, 58, 277–283, 1993.
19. M. Fliess, J. Lévine, P. Martin, and P. Rouchon, Flatness and defect of nonlinear systems: Introductory theory and examples, *Int. J. Control*, 6, 1327–1361, 1995.
20. M. Fliess, J. Lévine, P. Martin, and P. Rouchon, A Lie-Bäcklund approach to equivalence and flatness of nonlinear systems, *IEEE Trans. Autom. Contr.*, 44, 922–937, 1999.
21. R.B. Gardner, *The method of Equivalence and its Applications*, CBMS Regional Conference Series in Applied Mathematics, CBMS 58, SIAM, 1989.
22. R.B. Gardner and W.F. Shadwick, Symmetry and the implementation of feedback linearization, *Syst. Contr. Lett.*, 15, 25–33, 1990.
23. R.B. Gardner and W.F. Shadwick, The GS algorithm for exact linearization to Brunovský normal form, *IEEE Trans. Autom. Contr.*, 37, 224–230, 1992.
24. R.B. Gardner, W.F. Shadwick, and G.R. Wilkens, Feedback equivalence and symmetries of Brunovský normal forms, *Contemp. Math.*, 97, 115–130, 1989.
25. R.B. Gardner, W.F. Shadwick, and G.R. Wilkens, A geometric isomorphism with applications to closed loop controls, *SIAM J. Control Optim.*, 27 (6), 1361–1368, 1989.
26. J.W. Grizzle and S.I. Marcus, The structure of nonlinear systems possessing symmetries, *IEEE Trans. Autom. Contr.*, 30, 248–258, 1985.
27. B. Hamzi, *Analyse et commande des systèmes non linéaires non commandables en première approximation dans le cadre de la théorie des bifurcations*, Ph.D. thesis, University of Paris XI-Orsay, France, 2001.
28. B. Hamzi, J.-P. Barbot, and W. Kang, Bifurcation and topology of equilibrium sets for nonlinear discrete-time control systems, in *Proceedings of the Nonlinear Control Systems Design Symposium (NOLCOS'98)*, 1998, pp. 35–38.
29. B. Hamzi, J.-P. Barbot, and W. Kang, Stabilization of nonlinear discrete-time control systems with uncontrollable linearization, in *Modern Applied Mathematics Techniques in Circuits, Systems and Control*, World Scientific and Engineering Society Press, 1999, pp. 278–283.
30. B. Hamzi, J.-P. Barbot, and W. Kang, Normal forms for discrete time parameterized systems with uncontrollable linearization, *Proceedings of the 38th IEEE Conference on Decision and Control*, vol. 2, 1999, pp. 2035–2038.
31. B. Hamzi, J.-P. Barbot, and W. Kang, Bifurcation for discrete-time parameterized systems with uncontrollable linearization, *Proceedings of the 38th IEEE Conference on Decision and Control*, vol. 1, 1999, pp. 684–688.
32. B. Hamzi, J.-P. Barbot, S. Monaco, and D. Normand-Cyrot, Normal forms versus Naimark-Sacker bifurcation control, invited paper to the *Nonlinear Control Systems Design Symposium (NOLCOS'01)*, 2001.

33. B. Hamzi, J.-P. Barbot, S. Monaco, and D. Normand-Cyrot, Nonlinear discrete-time control of systems with a Naimark-Sacker bifurcation, *Syst. Control Lett.*, 44, 245–258, 2001.
34. B. Hamzi and I.A. Tall, Normal forms for nonlinear discrete-time control systems, *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2, 1357–1361, 2003.
35. J. Hauser, S. Sastry, and P. Kokotović, Nonlinear control via approximate input-output linearization, *IEEE Trans. Autom. Control*, 37, 392–398, 1992.
36. L.R. Hunt and R. Su, Linear equivalents of nonlinear time varying systems, in *Proceedings of the MTNS, Santa Monica, CA*, 1981, pp. 119–123.
37. L.R. Hunt, R. Su and G. Meyer, Design for multi-input nonlinear systems, in *Differential Geometric Control Theory*, R.W. Brockett, R. Millman, and H.J. Sussmann, Eds., Birkhauser, 1983, pp. 268–298.
38. A. Isidori, *Nonlinear Control Systems*, 3rd ed., Springer-Verlag, London, 1995.
39. M. Jankovic, R. Sepulchre, and P. Kokotovic, Constructive Lyapunov stabilization of nonlinear cascade systems, *IEEE Trans. Autom. Control*, 41, 1723–1735, 1996.
40. B. Jakubczyk, Equivalence and invariants of nonlinear control systems, in *Nonlinear Controllability and Optimal Control*, H.J. Sussmann, Ed., Marcel Dekker, New York, Basel, 1990, pp. 177–218.
41. B. Jakubczyk, Invariants of dynamic feedback and free systems, in *Proceedings of the European Control Conference, Groningen*, 1993.
42. B. Jakubczyk, Critical Hamiltonians and feedback invariants, in *Geometry of Feedback and Optimal Control*, B. Jakubczyk and W. Respondek, Eds., Marcel Dekker, New York, Basel, 1998, pp. 219–256.
43. B. Jakubczyk, Symmetries of nonlinear control systems and their symbols, *Can. Math. Conf. Proceed.*, 25, 183–198, 1998.
44. B. Jakubczyk, Feedback invariants and critical trajectories; Hamiltonian formalism for feedback equivalence, in *Nonlinear Control in the Year 2000*, vol. 1, A. Isidori, F. Lamnabhi-Lagarigue, and W. Respondek, Eds., LNCS vol. 258, Springer, London, 2000, pp. 545–568.
45. B. Jakubczyk and W. Respondek, On linearization of control systems, *Bull. Acad. Polon. Sci. Ser. Math.*, 28, 517–522, 1980.
46. B. Jakubczyk and W. Respondek, Feedback classification of analytic control systems in the plane, in: *Analysis of Controlled Dynamical Systems*, B. Bonnard, B. Bride, J.P. Gaúthier, I. Kupka, Eds., Birkhäuser, Boston, 1991, pp. 262–273.
47. V. Jurdjevic, Casimir elements and Optimal control, in B. Jakubczyk, W. Respondek, and T. Rzezuchowski, Eds., *Geometry in Nonlinear Control and Differential Inclusions*, No. 32, Banach Center Publications, Warsaw, 1995, pp. 261–275.
48. V. Jurdjevic, Optimal control problems on Lie groups: crossroads between geometry, and mechanics, in *Geometry of Feedback and Optimal Control*, B. Jakubczyk and W. Respondek, Eds., Marcel Dekker, New York, 1997, pp. 257–303.
49. T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
50. W. Kang, Extended controller form and invariants of nonlinear control systems with single input, *J. Math. Syst., Estimation Control*, 4, 1994, 1–25.

51. W. Kang, Quadratic normal forms of nonlinear control systems with uncontrollable linearization, *Proceedings of the 34th CDC*, New Orleans, 1995.
52. W. Kang, Bifurcation and normal form of nonlinear control systems — part I and part II, *SIAM J. Control Optim.*, 36, 193–212 and 213–232, 1998.
53. W. Kang, Normal form, invariants, and bifurcations of nonlinear control systems in the particle deflection plane, in *Dynamics, Bifurcations and Control*, F. Colonius and L. Grüne, Eds., LNCS Vol. 273, Springer, Berlin, Heidelberg, 2002, pp. 67–87.
54. W. Kang and A.J. Krener, Extended quadratic controller normal form and dynamic feedback linearization of nonlinear systems, *SIAM J. Control Optim.*, 30, 1319–1337, 1992.
55. W. Kang, Bifurcation control via state feedback for systems with a single uncontrollable mode, *SIAM J. Control Optim.*, 38, 1428–1452, 2000.
56. A.J. Krener, Approximate linearization by state feedback and coordinate change, *Syst. Control Lett.*, 5, 181–185, 1984.
57. A.J. Krener, S. Karahan, M. Hubbard, and R. Frezza, Higher order linear approximations to nonlinear control systems, *Proceedings of the 26th IEEE Conference on Decision and Control*, Los Angeles, 1987, pp. 519–523.
58. A.J. Krener, W. Kang, and D.E. Chang, Normal forms of linearly uncontrollable nonlinear control systems with a single input, NOLCOS, St. Petersburg, 2001.
59. A.J. Krener, W. Kang, and D.E. Chang, Normal forms and bifurcations of control systems, in *Proceedings of 39th IEEE Conference on Decision and Control*, 2000, pp. 1602–1607.
60. A.J. Krener, W. Kang, and D.E. Chang, Control Bifurcations, *IEEE Trans. Automa. Control*, 49 (8), 1231–1246, 2004.
61. A.J. Krener and L. Li, Normal forms and bifurcations of discrete time nonlinear control systems, *SIAM J. Control Optim.*, 40, 1697–1723, 2002.
62. I. Kupka, On feedback equivalence, in *Proceedings of the Canadian Mathematical Society Conference*, vol. 12, 1992, pp. 105–117.
63. A. Marigo, Constructive necessary and sufficient conditions for strict triangularizability of driftless nonholonomic systems, in *Proceedings of the 34th CDC*, Phoenix, Arizona, USA, 1999, pp. 2138–2143.
64. F. Mazenc and L. Praly, Adding integrations, saturated controls, and stabilization for feedforward forms, *IEEE Trans. Autom. Control*, 41, 1559–1578, 1996.
65. F. Mazenc and L. Praly, Asymptotic tracking of a reference state for systems with a feedforward structure, *Automatica*, 36, 2000, 179–187.
66. H. Nijmeijer and A.J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.
67. H. Poincaré, Sur les propriétés des fonctions définies par les équations aux différences partielles, *Oeuvres*, pp. XCIX–CX, Gauthier-Villars, Paris, 1929.
68. J.-B. Pomet, A differential geometric setting for dynamic equivalence and dynamic linearization, in *Geometry in Nonlinear Control and Differential Inclusions*, B. Jakubczyk, W. Respondek, and T. Rzezuchowski, Eds., vol. 32, Banach Center Publications, Warszawa, 1995, pp. 319–339.
69. W. Respondek, Feedback classification of nonlinear control systems in $\mathbb{R}^2$ and $\mathbb{R}^3$, in *Geometry of Feedback and Optimal Control*, B. Jakubczyk and W. Respondek, Eds., Marcel Dekker, New York, 1998, pp. 347–382.

70. W. Respondek, Symmetries and flat outputs of nonlinear control systems, dans *New Trends in Nonlinear Dynamics and Control*, W. Kang, M. Xiao, and C. Borges, Eds., LNCS vol. 295, Springer, 2003.
71. W. Respondek, Transforming a single input system to a p -normal form via feedback, in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, pp. 1574–1579.
72. W. Respondek and I.A. Tall, How many symmetries does admit a nonlinear single-input control system around an equilibrium? in *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, Florida, 2001, pp. 1795–1800.
73. W. Respondek and I.A. Tall, Nonlinearizable single-input control systems do not admit stationary symmetries, *Syst. Control Lett.*, 46, 1–16, 2002.
74. W. Respondek and I.A. Tall, Strict feedforward form and symmetries of nonlinear control systems, in *Proceedings of the 43rd IEEE Conference on Decision and Control*, Bahamas, 2004 pp. 1611–1616.
75. W. Respondek and M. Zhitomirskii, Feedback classification of nonlinear control systems on 3-manifolds, *Math. Control, Signals Syst.*, 8, 299–333, 1995.
76. R. Sepulchre, M. Janković, and P. Kokotović, *Constructive Nonlinear Control*, Springer, Berlin, 1996.
77. I.A. Tall, Classification des systèmes de contrôles non linéaires à une entrée, Thèse de 3^{ème} cycle, University of Dakar, 1999.
78. I.A. Tall and W. Respondek, Transforming nonlinear single-input control systems to normal forms via feedback, in *Proceedings of the Mathematical Theory of Networks and Systems*, Perpignan, 2000.
79. I.A. Tall and W. Respondek, Normal forms, canonical forms, and invariants of single input nonlinear systems under feedback, in *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, 2000, 1625–1630.
80. I.A. Tall and W. Respondek, Transforming a single-input nonlinear system to a feedforward form via feedback, in *Nonlinear Control in the Year 2000*, Vol. 2, A. Isidori, F. Lamnabhi-Lagarigue and W. Respondek, Eds., LNCS Vol. 259, Springer, London, 2000, pp. 527–542.
81. I.A. Tall and W. Respondek, Normal forms and invariants of nonlinear single-input systems with uncontrollable linearization, in *Proceedings of NOLCOS'2001*, ST. Petersburg, Russia, 2001, pp. 139–144.
82. I.A. Tall and W. Respondek, Feedback equivalence to feedforward forms for nonlinear single-input control systems, in *Dynamics, Bifurcations and Control*, F. Colonius and L. Grüne, Eds., LNCS vol. 273, Springer, Berlin, 2002, pp. 269–286.
83. I.A. Tall and W. Respondek, Feedback classification of nonlinear single-input control systems with controllable linearization: normal forms, canonical forms, and invariants, *SIAM J. Contr. Optim.*, 41, 1498–1531, 2003.
84. I.A. Tall and W. Respondek, Feedback equivalence to a strict feedforward form of nonlinear single-input control systems, *Int. J. Control*, in press.
85. I.A. Tall and W. Respondek, Feedback classification of nonlinear single-input control systems with uncontrollable linearization, (submitted).
86. I. A. Tall and W. Respondek, Normal forms of two-inputs nonlinear control systems, submitted in *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, USA, 2002, pp. 2732–2737.

87. I.A. Tall and W. Respondek, Smooth Analytic normal forms: a special class of strict feedforward systems, in *Proceedings of Symposium on Nonlinear Control and Observe Design, SYNCOD*, Stuttgart, Springer, 2005.
88. I.A. Tall, Normal forms of multi-inputs nonlinear control systems with controllable linearization, in *New Trends in Nonlinear Dynamics and Control, and their applications*, W. Kang, M. Xiao, and C. Borges, Eds., LNCS vol. 295, Springer, Berlin-Heidelberg, 2003, pp. 87–100.
89. I.A. Tall, Feedback classification of multi-input nonlinear control systems, *Siam J. Control Optim.*, 43, 2049–2070, 2005.
90. A. Teel, *Feedback stabilization: nonlinear solutions to inherently nonlinear problems*, Memorandum UCB/ERL M92/65.
91. A. Teel, A nonlinear small gain theorem for the analysis of control systems with saturation, *IEEE Trans. Autom. Control*, 41, 1256–1270, 1996.
92. S. Sastry, *Nonlinear Systems: Analysis, Stability, and Control*, Springer-Verlag, 1999.
93. H.J. Sussmann, Symmetries and integrals of motion, in *Geometry in Nonlinear Control and Differential Inclusions*, *ibid.* pp. 379–393.
94. A.J. van der Schaft, Symmetries in optimal control, *SIAM J. Control Optim.*, 25, 245–259, 1987.
95. M. Zhitomirskii and W. Respondek, *Simple Germs of Corank One Affine Distributions*, *Banach Center Publications*, vol. 44, 1998, pp. 269–276.

5

Singular Perturbation and Chaos

M. Djemai and S. Ramdani

CONTENTS

5.1	Introduction	263
5.2	Autonomous Chaotic Dynamical Systems	264
5.2.1	Background on Nonlinear Dynamical Systems	264
5.2.2	Linearization, Stability, and Invariant Manifolds	265
5.2.3	Strange Attractors	265
5.2.4	Structural Stability and Bifurcations	266
5.3	Singularly Perturbed Systems	267
5.3.1	Some Fundamental Notions	268
5.3.2	Integral Manifold Approach	272
5.3.2.1	Approached Resolution of the Condition (5.22)	274
5.4	Illustrative Examples	275
5.4.1	An Electronic Oscillator: Chua's Cubic System	276
5.4.2	A Neuron Model: The Hindmarsh and Rose System	280
	References	285

5.1 Introduction

In the last decade, significant advances have been made in the theory of nonlinear state feedback control [3]. It is well known that many physical systems naturally possess a time-scale separation. Singular perturbation theory provides the mean to decompose such systems into slow and fast dynamics which greatly simplifies their structural analysis and any subsequent control design. The most significant development in the analysis and control of nonlinear *singularly perturbed systems* has been the *integral manifold approach* [5, 6, 34]. The dynamical behavior of such systems is geometrically captured by the rapid approach of the fast system states to the attractive slow integral manifold. The fundamental property of the

integral manifold is that once the fast system states hit the integral manifold, they remain on the manifold thereafter. This fundamental property is described by the integral manifold condition, an ε -dependent partial differential equation. The small parameter ε is the speed ratio of the slow versus the fast system dynamics [2].

For nonlinear singularly perturbed systems, the composite state feedback law is the sum of slow and fast parts. The fast one steers the fast system states to the slow integral manifold. It is designed so as to be inactive on this manifold. By equating coefficients to like powers of ε in the integral manifold condition, it becomes possible to design a slow control that will steer the system along a desired manifold to within any required order of ε accuracy.

The paradigm of deterministic chaos is certainly one of the most interesting phenomena observed and studied during the second half of the 20th century. It has changed our vision of some irregular and apparently stochastic behaviors of systems involved in many fields of modern science, from biology to celestial mechanics. There are many different approaches in the study of chaotic dynamical systems. We can distinguish two large fields in the analysis of these problems: a theoretical one based on mathematical analysis and an experimental one using numerical methods. This second more empirical field is actually developed through nonlinear time series analysis methods (see, e.g., [20]).

In this chapter, we will briefly recall some of the basic tools of the mathematical and geometrical theories of dynamical systems (which are detailed in Chapter 2). In particular, we will focus on those tools characterized by chaos and singular perturbation phenomenon.

5.2 Autonomous Chaotic Dynamical Systems

5.2.1 Background on Nonlinear Dynamical Systems

From a mathematical point of view, chaotic dynamical systems are deterministic and nonlinear systems which exhibit the property of sensitive dependence on initial conditions (SDIC). This dependence means that two initially very close states of the system will exponentially diverge over the course of time. This divergence is quantified by the Lyapunov exponent.

If time is considered as a continuous variable, a chaotic dynamical system is often defined by a set of first-order ordinary differential equations (ODE):

$$\frac{dx}{dt} = \dot{x} = f(x) \quad (5.1)$$

where $t \in I \subset \mathbb{R}$, $x \in \mathbb{R}^n$ (n -dimensional space), and $f : U \rightarrow \mathbb{R}^n$ is a smooth function defined on some subset. If the vector field f does not contain time explicitly, the system defined by (5.1) is said to be autonomous. A solution or flow of (5.1) defined by $\phi : U \times I \rightarrow \mathbb{R}^n$, $\phi : (x, t) \mapsto \phi(x, t)$ satisfies:

$$\left[\frac{d}{dt} \phi(x, t) \right]_{t=s} = f(\phi(x, s)) \quad (5.2)$$

for all $x \in U$ and $s \in I$. See Chapter 2 for more details and theoretical results.

5.2.2 Linearization, Stability, and Invariant Manifolds

The description of the solutions of (5.1) near a fixed point $\bar{x}$ can be made using a local linearization and by studying the linear system:

$$\dot{\delta} = J_f(\bar{x})\delta \quad (5.3)$$

where $J_f(\bar{x}) = (\partial f_i / \partial x_j)_{x=\bar{x}}$ is the Jacobian matrix of f at $\bar{x}$ and $\delta = x - \bar{x} \in \mathbb{R}^n$, $\|\delta\| \ll 1$ (where $\|\cdot\|$ denotes the Euclidean norm).

Linear systems such as (5.3) are well known (Chapter 2 of [17]) and an integration leads to the solution flow $e^{tJ_f(\bar{x})}$.

5.2.3 Strange Attractors

In this chapter, we are interested in dissipative systems, which are the most common ones in physics and engineering. The dissipative property implies a constraint on the system's dynamics.

From a physical point of view, dissipative systems do not satisfy energy conservation laws [10]. Mathematically, this means that these systems do not have a Hamiltonian independent of time.

In the phase space, dissipative systems are characterized by an attraction of all trajectories to a geometrical object called the attractor. This subset of the phase space is invariant under the transformation of the flow defined by Equation (5.1) over the course of time. There are many definitions of an attractor in the literature [1, 15] (see also Chapter 2) which are not always equivalent. The following definition is proposed in [1]. First we have to define the absorbing domains of the phase space associated to the solution flows of Equation (5.1):

DEFINITION 1 A domain of the phase space is said to be absorbing if it completely contains all positive semi-trajectories starting in it. If, in

addition, every solution orbit falls into it after a finite time, it is said to be globally absorbing.

We now come to the definition of an attractor defined by a solution flow ϕ of Equation (5.1):

DEFINITION 2 If B is an absorbing domain with compact closure $\bar{B}$, then the set defined by:

$$A = \bigcap_{t>0} \phi(\bar{B}, t)$$

is called an *attractor* of Equation (5.1). The attractor is said to be maximal if B is globally absorbing.

Some attractors are very simple, such as stable fixed points for systems evolving toward stationary states or limit cycles that are the characterization of periodic dynamics. Attractors that are not points or cycles are called strange attractors. This term was first introduced by Ruelle and Takens [31] in their famous paper describing the nature of turbulence. The Lorenz attractor [15, 23] or Chua's attractor [11, 12] are some of the most studied ones.

For a chaotic system, the geometry of the attractor can be very complicated. This kind of attractor, defined by autonomous systems of ODEs given by Equation (5.1), cannot appear in phase spaces of less than three dimensions. They are characterized by noninteger fractal dimensions, and some of them show a self-similarity property which means that their global geometrical structures can be observed at different scales. One can find more details in the work of Mandelbrot, a pioneer in the field of fractal phenomena [24].

5.2.4 Structural Stability and Bifurcations

Another important concept of the theory of dynamical systems is the occurrence of bifurcations. Many dynamical systems have parameters appearing in their equations. When these parameter values (bifurcation values) are changed, one may observe modifications (bifurcations) of the qualitative structure of their solution flows. The local bifurcations of a system are studied by analyzing the vector field near an equilibrium point [9] (also see Chapter 2). Global bifurcations are related to the description of the global changes of flows when local analysis is not useful. In the global approach, particular trajectories are studied like homoclinic orbits (orbits connecting fixed points to themselves) or heteroclinic orbits (orbits connecting distinct fixed points).

5.3 Singularly Perturbed Systems

The notion of singular perturbation disturbance in the analysis of dynamical systems is very important and represents the theoretical base of numerous modern concepts (i.e., bifurcation, chaos, regular perturbation, singular perturbation, etc.) This last case was originally developed for the analysis of phenomena with multiple time-scales characterizing the evolution of systems belonging to the domain of the mechanics of fluids. Later, this theory was extended to the study of the other phenomena characterized by singular perturbation. For example, induction motor [35, 36], robotics [39], etc.

In 1952, Tikhonov discussed conditions for which the solution of a system of strangely disrupted common differential equations could be approached asymptotically by the solutions of two sub-systems: *slow* and *fast* [45]. In 1963, Vasiléva [46] proposed a solution to resolve ODE with singular perturbation [22] in terms of a sum of three series of solutions: *internal*, *external*, and *intermediary*.¹

The introduction of the singular perturbation theory in the field of the control theory [40] led to results concerning the control of the singularly perturbed nonlinear systems from the concept of the well-known *composed control* [4, 41–43], which give a panorama of different and varied applications developed in this domain. The application of these results is relevant to two big problems of analysis in theory of control:

1. Systems of big dimension, implying “heavy” control to implement
2. Combined presence of physical phenomena in several scales of time, implying important numeric difficulties

Recently, another approach, *integral manifold* was introduced to study, in a geometrical way, the behavior of the solutions of a singularly perturbed system [4, 37]. According to this approach, the fast dynamics of the system are forced on an attractive manifold on which the evolution of the system is described with a reduced model (also called *slow system*). The slow system has to represent, after a small period of time, the dynamics of the real system. The objectives of control can be fixed on the basis of this model. This simplifies the control procedure. Several applications using this approach were developed, notably in the domains of the robotics [44] and electrical machines [36].

¹Internal solution (resp. external) describes the behavior of the system in (resp. outside) the limit layer; the *intermediary* solution connects the two other solutions.

5.3.1 Some Fundamental Notions

Consider the following perturbed system:

$$\dot{\chi} = f(\chi, \varepsilon) \quad (5.4)$$

with $\chi \in \mathbb{R}^n$ and $\varepsilon \in V_0 \subset \mathbb{R}$.

In Reference [5], the authors present a determining criterion for physical process of the form (5.4) to be written in the following *standard form*:

$$\dot{x} = f(x, z, \varepsilon, t) \quad x(t_0) = x_0 \quad x \in \mathbb{R}^n \quad (5.5)$$

$$\varepsilon \dot{z} = g(x, z, \varepsilon, t) \quad z(t_0) = z_0 \quad z \in \mathbb{R}^m \quad (5.6)$$

Functions f and g are supposed to be enough continuously differentiable with respect to their arguments. Parameter ε is supposed to be “small” and represent the speed of evolution of the state dynamics z , (*fast dynamics*), with respect to x (*slow dynamics*).

The classic method of analysis of singularly perturbed systems consist in showing a reduced order model obtained by putting $\varepsilon = 0$ in model (5.5) and model (5.6). Dimension of the state space then falls from $n + m$ to n , because differential equation (5.6) “degenerates” into one algebraic equation:

$$g(\bar{x}, \bar{z}, 0, t) = 0 \quad (5.7)$$

where $\bar{x}, \bar{z}$ indicate the variable of system (5.5) and system (5.6) for $\varepsilon = 0$.

System (5.5) and system (5.6) are in *standard form* if and only if the following hypothesis (5.7) is verified:

HYPOTHESIS 1

*In the considered domain, (5.7) possesses real roots isolated:*²

$$\bar{z} = \Phi(\bar{x}, t)$$

Hypothesis 1 holds when the function $g(x, z, 0, t)$ is such that $\partial g / \partial z$ is regular; the solution is then unique. Furthermore, this allows us to use the implicit function theorem:

THEOREM 1 (Implicit function theorem)

Let us consider two sets $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$ and one function of class C^∞ , $F : A \times B \rightarrow \mathbb{R}^n$. One notes $(x, y) = (x_1, \dots, x_m, y_1, \dots, y_n)$ a point of $A \times B$. Suppose that for a certain point $(x^0, y^0) \in A \times B$:

$$F(x^0, y^0) = 0$$

²The root $\Phi(\bar{x}, t)$ is *isolated* if there exists $\varepsilon > 0$ such as Equation (5.7) does not admit another solution different from this one for $|\bar{z} - \Phi(\bar{x}, t)| < \varepsilon$.

and that the matrix:

$$\frac{\partial F}{\partial y} = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n} \end{bmatrix}$$

is not singular at the point (x^0, y^0) . Then, there are two open neighborhoods A_0 of x^0 in A and B_0 of y^0 in B as well as one unique function of class C^∞ , $G : A_0 \rightarrow B_0$, such that:

$$F(x, Gx) = 0$$

for any $x \in A_0$, that is, $y = G(x)$ is a solution of $F(x, y) = 0$ with respect to y , defined around the point (x^0, y^0) .

To resolve Equation (5.7): solution $\bar{z}$ is an analytical function of the variable t and x . By using Gröbner's formula³ [38]; this solution can be given by the following expression:

$$\bar{z} = \{e^{-\sum_{i=1}^n g_i L_{\gamma_i}}(I_d)\}_{|z=z^0}$$

where g_i is the i th component of the vector g and

$$\gamma_i = \left\{ \left[\frac{\partial g}{\partial z} \right]_{|z=z^0} \right\}_i$$

THEOREM 2 [38]

Inverse function ϕ , solution of the equation,

$$Y = h(w) \quad \text{with} \quad Y \in \mathbb{R}^n, \quad \text{and} \quad w \in \mathbb{R}^n$$

is given by the following formula

$$w = \phi(y, w_0) = \{e^{F(y, w_0, \cdot)} I_d\}_{|w=w_0}$$

$$F(y, w_0, \cdot) = \sum_{i=1}^k (y^i - h(w_0)^i) L_{D_i}(\cdot)$$

³Gröbner's formula allows the calculation of the solution of an analytic equation of the form:

$$Y = h(w) \quad \text{with} \quad Y \in \mathbb{R}^n \quad \text{and} \quad w \in \mathbb{R}^n$$

where h is an analytic function. One supposes that $(\partial h / \partial w)_{w=w_0}$ is invertible for any w_0 for the neighborhood of it solution.

where D_i represent the i^{th} column of the inverse matrix it Jacobian of the function:

$$D \stackrel{\text{def}}{=} \left(\frac{\partial h}{\partial w} \right)_{w=w_0}^{-1}$$

and y^i and h^i represent, respectively, the i^{th} component of vectors y and h .

Hypothesis 1 assures that for every root verifying Equation (5.7) one can associate a reduced model of order n correctly defined. To obtain the reduced model, one substitutes Φ in $\bar{z}$ in (5.5) to obtain:

$$\dot{\bar{x}} = f(\bar{x}, \Phi(\bar{x}, t), 0, t), \quad \bar{x}(T_0) = x_0 \quad (5.8)$$

Later, one considers this model expressed in the compact form by considering implicitly the function Φ , or:

$$\dot{\bar{x}} = f(\bar{x}, 0, t), \quad \bar{x}(t_0) = x_0 \quad (5.9)$$

Model (5.9) is called a *slow model*, (almost *steady-state model*).

Consider the following question: Given $T > t_0$, are the following approximations valid uniformly on the interval of time $[t_0, T]$?

$$x(t) = \bar{x}(t) + O(\varepsilon) \quad (5.10)$$

$$z(t) = \bar{z}(t) + O(\varepsilon) \quad (5.11)$$

Solution $\bar{z}$ satisfies algebraic Equation (5.7), and $\bar{z}(t_0)$ is generally different from the given initial condition z_0 ; approximation (5.11) is not so valid when t is equal (or close) to t_0 . Consequently, (5.11) is valid only outside the *limit layer* defined for any $t \in [t_0, t_1]$. Approximation (5.10) is uniformly valid because we can always choose x_0 as the initial condition for variable $\bar{x}$ (with $\bar{x}(t_0) = x_0$).

To analyze the behavior of the solution $\bar{z}$ in the limit layer [viz, where the approximation (5.11) is not valid], one introduces the change of time-scale:

$$\tau = \frac{t - t_0}{\varepsilon}$$

In this scale of “dilated” time, system (5.5) and system (5.6) are written, respectively, as:

$$\frac{dx}{d\tau} = \varepsilon f(x, z, \varepsilon, \varepsilon\tau + t_0) \quad (5.12)$$

$$\frac{dz}{d\tau} = g(x, z, \varepsilon, \varepsilon\tau + t_0) \quad (5.13)$$

Equation (5.12) and Equation (5.13) show that z evolves more quickly than x which stays in its limit layer, close to its initial values $x(0) = x_0$.

By assuming $\varepsilon = 0$ in (5.12) and (5.13), one obtains the fast dynamics, described by the equation:

$$\frac{dz}{d\tau} = g(x_0, z, 0, t_0) \quad (5.14)$$

To determine the component of the solution $z(t)$ which evolves quickly in the limit layer, one defines the variable $\hat{z} = z - \bar{z}$.

Because $d\bar{z}/d\tau = 0$ [due to the fact that x is supposed to be constant in the fast time-scale and when $\bar{z} = \Phi(x, \varepsilon)$], the evolution of $\hat{z}$ in the limit layer is described by the following dynamics:

$$\frac{d\hat{z}}{d\tau} = g(x_0, \hat{z}(\tau) + \bar{z}(t_0), 0, t_0) \quad (5.15)$$

with $\hat{z}(0) = z_0 - \bar{z}(t_0)$.

Therefore, a “finer” approximation [with regard to that given in (5.11)] of the exact solution $z(t)$ is given by the following relation:

$$z(t) = \bar{z}(t) + \hat{z}\left(\frac{t - t_0}{\varepsilon}\right) + O(\varepsilon) \quad (5.16)$$

REMARK 1

To improve the approximation order in ε , we can substitute the function $\Phi(x, \varepsilon)$ at any prefixed order of approximation to z in (5.5).

REMARK 2

With regard to (5.11), approximation (5.16) contains the term $\hat{z}(\tau)$ which represents the fast component of the solution z , whereas $\bar{z}$ represents the slow component. In fact:

$$\frac{d\hat{z}}{dt} = \frac{d\hat{z}}{d\tau} \frac{d\tau}{dt} = \frac{1}{\varepsilon} \frac{d\hat{z}}{d\tau} = \frac{1}{\varepsilon} g(x_0, \hat{z}(\tau) + \bar{z}(t_0), 0, t_0)$$

To guarantee the validity of approximation (5.16), one needs to study the stability of system (5.15). For this, the following hypotheses are proposed.

HYPOTHESIS 2

The appropriate values of the matrix $\{\partial g / \partial z\}$ estimated along $\bar{x}, \bar{z}$, for any $t \in [t_0, T]$, are with negative real parts.

If $\hat{z}(0)$ is not close to 0 [z_0 is very different of $\bar{z}(t_0)$], one needs a supplementary hypothesis to assure the stability of system (5.15).

HYPOTHESIS 3

The equilibrium point $\hat{z}(\tau) = 0$ of system (5.15) is uniformly asymptotically stable, for any x_0 and t_0 , and the initial point $\hat{z}(0) = z_0 - \bar{z}(t_0)$ belongs to the attraction domain of the manifold $z = \bar{z}$.

If Hypothesis 2 is satisfied, then $\lim_{\tau \rightarrow \infty} \hat{z}(\tau) = 0$ when $\tau \rightarrow \infty$, namely, solution $z(t)$ will be close to $\bar{z}$ from a certain moment $t_1 > t_0$.

It is important to note that if Hypothesis 3 is satisfied, then the roots of the algebraic Equation (5.7) are *isolated* and consequently Hypothesis 1 is also satisfied.

Finally, one recalls the fundamental theorem of Tikhonov, which allows to confirm approximations (5.10) and (5.16).

THEOREM 3 (Tikhonov [45])

If Hypotheses 2 and 3 are satisfied, then approximations (5.10) and (5.16) are valid for any $t \in [t_0, T]$ and there exists $t_1 \geq t_0$ such as (5.11) is valid for any $t \in [t_1, T]$.

REMARK 3

In Theorem 3, time instant t_1 depends strongly on the value of the parameter ε .

The reader is referred to Kokotovic et al. [4], where some variants of Theorem 3 are presented and applied.

5.3.2 Integral Manifold Approach

The main objective of the *integral manifold approach* is to present a geometrical point of view regarding the behavior of slow and fast states as trajectories in $\mathbb{R}^{n+m}$. One considers a class of singularly perturbed systems:

$$\dot{x} = f(x, z) \quad x(t_0) = x_0 \quad x \in \mathbb{R}^n \quad (5.17)$$

$$\varepsilon \dot{z} = g(x, z) \quad z(t_0) = z_0 \quad z \in \mathbb{R}^m \quad (5.18)$$

where the dependence of f and g in ε and t was removed so as to simplify the statement of the concepts which follow.

Let M_ε be the manifold of $\mathbb{R}^m$, parameterized by ε , and defined as:

$$M_\varepsilon \stackrel{\text{def}}{=} \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m : z = \Phi(x, \varepsilon)\} \quad (5.19)$$

DEFINITION 3 One says that M_ε is an invariant manifold for system (5.5) and system (5.6) if it is invariant with respect to the dynamics of this system, that is:

$$\begin{aligned} z(t_0, \varepsilon) &= \Phi(x(t_0, \varepsilon), \varepsilon) \\ \Rightarrow z(t, \varepsilon) &= \Phi(x(t, \varepsilon), \varepsilon) \quad \forall t \geq t_0 \end{aligned} \quad (5.20)$$

To obtain the manifold, one proceeds as follows:
By deriving (5.20) with respect to t , one obtains:

$$\dot{z} = \left(\frac{\partial \Phi}{\partial x} \right) f(x, \Phi(x, \varepsilon)) \quad (5.21)$$

by multiplying (5.21) by ε and by using (5.18), one obtains:

$$g(x, \Phi(x, \varepsilon), \varepsilon) \stackrel{\text{def}}{=} \varepsilon \left(\frac{\partial \Phi}{\partial x} \right) f(x, \Phi(x, \varepsilon)) \quad (5.22)$$

NOTE Relation (5.22), called *condition of invariant manifold*, must be satisfied by Φ for any x belonging to the considered domain, for any $t \geq t_0$, and for every $\varepsilon \in [0, \varepsilon^*]$, where ε is a constant positive. It characterizes the invariance of the variety M_ε .

The resolution of (5.22) with regard to Φ is difficult because it requires the resolution of a partial differential equation. As one will see later, the development of the exact solution $\Phi(., \varepsilon)$, allows one to find approached solutions (explicit) which can be satisfactory for applications (recall that parameter ε is small).

η , the distance between the state z and its manifold M_ε , is:

$$\eta \stackrel{\text{def}}{=} z - \Phi(x, \varepsilon)$$

Then system (5.17) and system (5.18) expressed as a function of x and, from η , respectively, we obtain

$$\begin{aligned} \dot{x} &= f(x, \Phi(x, \varepsilon) + \eta) \\ \varepsilon \dot{\eta} &= g(x, \Phi(x, \varepsilon) + \eta) - \varepsilon \left(\frac{\partial \Phi}{\partial x} \right) f(x, \Phi(x, \varepsilon) + \eta) \end{aligned} \quad (5.23)$$

The invariance of the manifold M_ε is then characterized by $\eta = 0$. Indeed, if $\eta = 0$ (i.e., if one is on the manifold M_ε), then one obtains $\dot{\eta} = 0$.

Consequently, the evolution of system (5.17) and system (5.18) can be expressed by the dynamics of the following reduced system:

$$\dot{x} = f(x, \Phi(x, \varepsilon)) \quad \text{with } x(t_0) = x_0 \quad (5.24)$$

Then, one calls M_ε a *slow manifold*, in the sense that in every manifold M_ε corresponds one *slow model* of (5.24).

When $\eta \neq 0$, by putting formally $\varepsilon = 0$ in (5.23) and (5.24) and by considering the change of time-scale $\tau = (t - t_0)/\varepsilon$, the equation expressing the dynamics of the variable $\eta(\tau)$ in the new time-scale, is given by:

$$\frac{d\eta}{d\tau} = g(x_0, \Phi(x_0, 0) + \eta(\tau)) \quad (5.25)$$

REMARK 4

Equation (5.25) is similar to Equation (5.15); consequently, the variable $\hat{z}$, corresponds to the variable η in the approach of integral manifold and $\bar{z}$ corresponds to $\Phi(x, 0)$.

Solution $z(t)$ is then written as

$$z(t) = \Phi(x, 0) + \eta(\tau) + O(\varepsilon)$$

Under Hypotheses 2 and 3 of Theorem 3, M_0 is a stable manifold (attractive) of the system (5.25). This system describes trajectories of x and η , which, for any given x_0 , are on one *fast manifold* defined by $x = x_0 = \text{constant}$, and approaches quickly toward the manifold M_0 . For $\varepsilon \neq 0$ but "small," fast manifold shows solutions which quickly approach the slow manifold M_ε .

5.3.2.1 Approached Resolution of the Condition (5.22)

This procedure requires that functions f and g can be developed in a Taylor's series with power of ε . For function g , one obtains:

$$g(x, \Phi_0(x) + \varepsilon \Phi_1(x) + \dots) = g(x, \Phi_0(x)) + \frac{\partial g}{\partial z} \varepsilon \Phi_1(x) + \dots \quad (5.26)$$

where $\partial g / \partial z$ and all the derivatives of higher order are estimated in x and $z = \Phi_0(x)$. One calculates the development of the function $f(x, \Phi, \varepsilon, t)$ around the point $\varepsilon = 0$.

By substituting function Φ in the condition (5.22) and by the development of a Taylor's series with powers of ε :

$$\Phi(x, \varepsilon) = \Phi_0(x) + \varepsilon \Phi_1(x) + \varepsilon^2 \Phi_2(x) + \dots \quad (5.27)$$

one can calculate iteratively Φ_0, Φ_1 , etc., and then identify the terms of the same power in ε .

As an immediate consequence of (5.22), (5.19), (5.27), and (5.26), one obtains:

$$g(x, \Phi_0(x), u) = 0 \quad (5.28)$$

$$\Phi(x, 0) = \Phi_0(x) \quad (5.29)$$

REMARK 5

One notes that solution Φ_0 corresponds to solution $\bar{z}$ in the algebraic Equation (5.7).

With $\Phi_0(x)$ known, one identifies terms in ε of power 1, and one obtains:

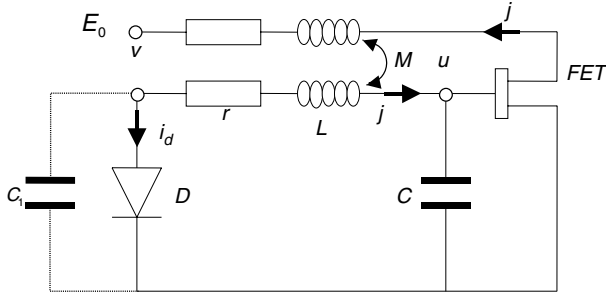
$$\Phi_1(x) = \left[\frac{\partial g}{\partial z} \right]_{z=\Phi_0(x)}^{-1} \frac{\partial \Phi_0(x)}{\partial x} f(x, \Phi_0(x))$$

Now, with $\Phi_0(x)$ and $\Phi_1(x)$ known, one can identify terms in ε of power 2, and obtain $\Phi_2(x)$ and so on until some wished order [4].

As the existence of solutions $\Phi_1, \Phi_2, \dots$ is connected to the invertibility of the Jacobian of the function g , evaluated at $z = \Phi_0$, this invertibility is guaranteed by Hypothesis 1.

5.4 Illustrative Examples

In this section, we describe two singularly perturbed chaotic systems without applying all the previous theoretical results. Our purpose is to show that this type of model can be found in different scientific fields (e.g., electronics and neurophysiology) and to point out some of the advantages of the singular perturbation methods. For both these systems, we will compute the slow manifold M_0 and describe some of their dynamical properties. One of the main ideas of the singular perturbation approach in the study of chaotic dynamical systems is to provide geometrical elements for the description of these systems. These elements can help in understanding

**FIGURE 5.1**

The Chua's cubic electronic oscillator.

and controlling their chaotic dynamics since they are based on analytical equations that are not sensitive to initial conditions.

5.4.1 An Electronic Oscillator: Chua's Cubic System

The following model describes the dynamics of an electronic relaxation oscillator circuit with a nonlinear cubic characteristic [11, 28]. It is composed of a harmonic oscillator based on a field effect transistor (FET) coupled to a relaxation oscillator based on a tunnel diode (see Figure 5.1).

The electric analysis of this circuit leads to the following three-dimensional nonlinear system:

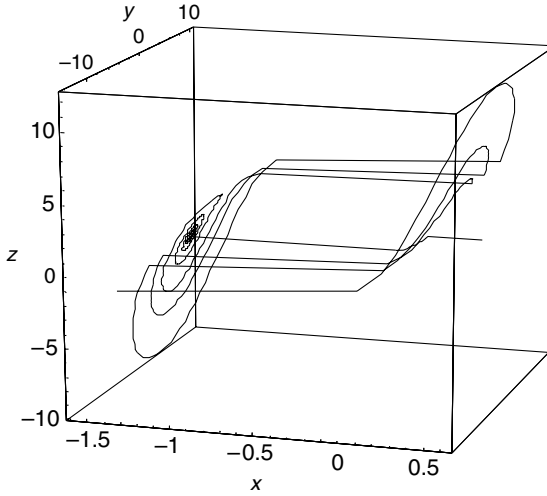
$$\begin{aligned}\varepsilon \dot{x} &= z - F_\mu(x) \\ \dot{y} &= -z \\ \dot{z} &= cx + y + dz\end{aligned}\tag{5.30}$$

where x, y , and z are, respectively, proportional to the voltages v, u and to the current i . $\varepsilon = C_1/C$ ($C_1 \ll C$), $F_\mu(x) = ax^3 + bx^2 + \mu x$ which is due to the shape of the i_d function of v , and μ is a bifurcation parameter; c and d are related to the electrical components of the circuit r, L, C, M . Generally, the parameters values are: $\varepsilon = 0.001, a = 44/3, b = 41/2, c = -0.7$, and $d = 0.24$. The fast system associated to the slow one (5.30) is defined by

$$\begin{aligned}x' &= z - F_\mu(x) \\ y' &= -\varepsilon z \\ z' &= \varepsilon (cx + y + dz)\end{aligned}\tag{5.31}$$

where $x' = dx/d\tau$ and $t = \varepsilon\tau$.

For $\mu = 2$, we observe a chaotic attractor showed on Figure 5.2. This model has three fixed points: two unstable focuses (O_1 and O_2) and a

**FIGURE 5.2**

Chua's cubic attractor for $\mu = 2$ (5.30) with initial conditions: $x_0 = 0.5, y_0 = -0.5, z_0 = 1$.

saddle (P). The coordinates of these points can be found by writing the vanishing condition of the vector field defined by (5.30):

$$\begin{aligned} z - F_\mu(x) &= 0 \\ z &= 0 \\ cx + y + dz &= 0 \end{aligned} \tag{5.32}$$

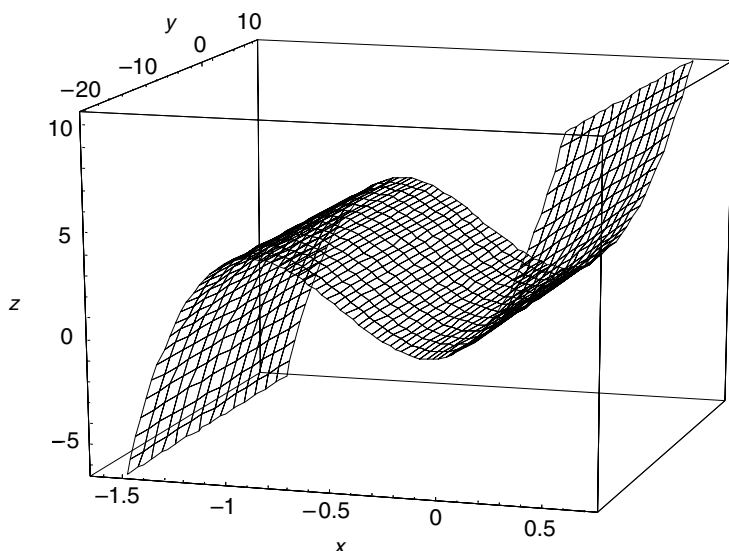
which implies:

$$\begin{aligned} F_\mu(x) &= 0 \\ y &= -cx \\ z &= 0 \end{aligned} \tag{5.33}$$

The resolution of the third-order equation $F_\mu(x) = 0$ for $\mu = 2$ leads to three real distinct solutions which are the abscises of the three fixed points of (5.30) in the phase space. The second equation of (5.33) gives the y -coordinate. Finally, the third equation of (5.33) states that these three points are situated in the plane $z = 0$.

Equations (5.30) define a singularly perturbed two time-scale model with one fast variable (x) and two slow ones (y and z). The slow or critical manifold [obtained by setting $\varepsilon = 0$ in (5.30), see [19, 21]) associated to this singularly perturbed system is a cubic cylinder defined by (see Figure 5.3):

$$M_0 = \left\{ (x, y, z) \in \mathbb{R}^3 : z = F_\mu(x) \right\} \tag{5.34}$$

**FIGURE 5.3**

Slow manifold M_0 associated to system (5.30).

This manifold is composed of two stable attractive parts on which the slow dynamics take place (see Figure 5.4). The fast dynamics, which are parallel to the x -axis, connect the two slow parts. The global dynamics are alternatively slow and fast following a sort of dichotomy of motion.

If an initial state is located near the first unstable focus O_1 , the trajectories will slowly move away around it until they reach an unstable part of M_0 , called the *fold* [7, 8]. There, the fast motion, which is parallel to the x -axis, replaces the slow one. Then, the trajectory takes place with a slow motion on a part of the second attractive sheet of M_0 called the *co-fold*. The co-fold is simply the projection of the fold along the direction of the fast motion (see Figure 5.4 and Figure 5.5).

After some time, this trajectory will reach the second fold of the other attractive part where the fast dynamics *re-inject* the solution near the initial conditions on the first attractive sheet in the neighborhood of the first unstable focus O_1 (see Figure 5.5). This re-injection process, which is essential in this type of phenomena, does not always take place at the same point of the neighborhood of the fixed point O_1 . This is due to the sensitivity to initial conditions: a small perturbation of these conditions will affect the re-injection point and the number of divergent oscillations observed on the first attractive sheet of the slow manifold M_0 .

Note that contrary to the first unstable focus O_1 , the saddle point P and the second unstable focus O_2 are located on folds (see Figure 5.5).

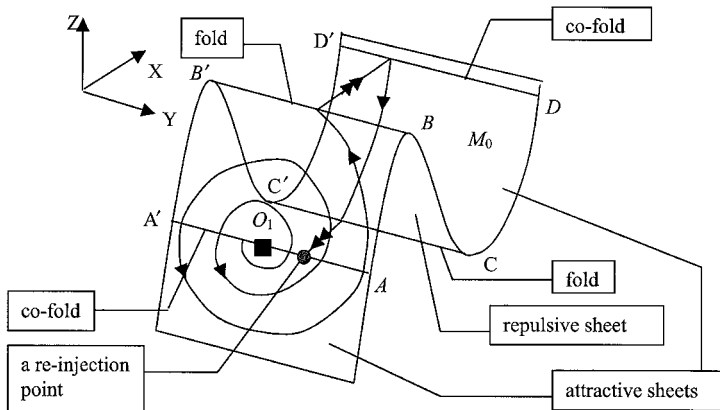


FIGURE 5.4

A representation of the slow manifold M_0 of Chua's cubic system. The fast trajectories, which are parallel to the x-axis, are drawn using double arrows. The slow ones are represented with simple arrows. The straight lines AA' and DD' indicate the co-folds and the lines BB' and CC' the folds. The co-folds are simply the projections of the folds along the direction of the fast motion. A re-injection point is also shown.

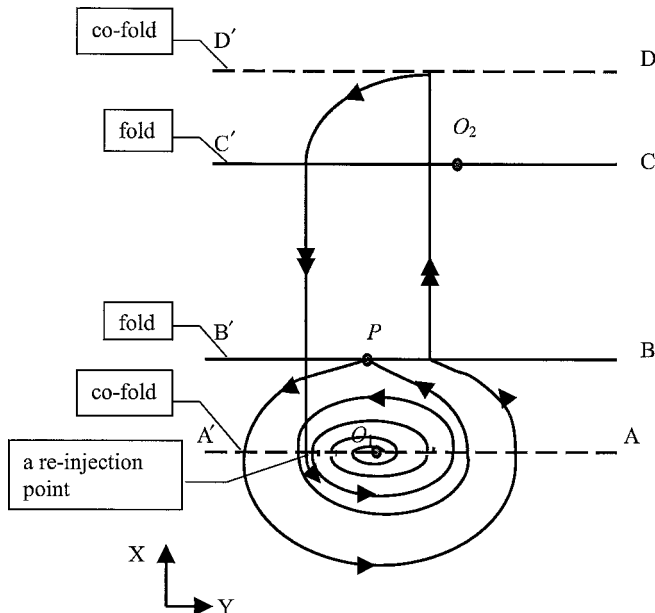


FIGURE 5.5

The global motion of the Chua's cubic system with the fast trajectories (double arrows) and the slow ones (simple arrows) is shown in the plane Oxy. The unstable focuses O_1 and O_2 are shown as the saddle point P . A re-injection point is also shown. The fixed points O_2 and P are located on folds and are called pseudo-singular points.

This destroys the symmetry of the dynamics on the two attractive sheets of M_0 . P and O_2 points are called *pseudo-singular points* (more details and alternative methods to study this type of models can be found in [7, 8, 28–30]).

Figure 5.2 is obtained by the numerical integration of Equations (5.30) with initial conditions $x_0 = 0.5$, $y_0 = -0.5$, and $z_0 = 1$.

5.4.2 A Neuron Model: The Hindmarsh and Rose System

The following model, the Hindmarsh and Rose model, (HR) is an extension of the FitzHugh–Nagumo system [14, 26] which is a reduced form of the famous Hodgkin–Huxley model [18]. It was proposed in 1984 [16] to describe the dynamics of action potentials in the neuron of a pond snail. It became a reference model in the nonlinear dynamical approach of studying neural systems behaviors [27]. Many different studies have been devoted to the HR model. For example, Wang proposed a mathematical analysis of the bursting oscillations observed in this model [33]. A bifurcation analysis was performed by Fan and Holden [13]. Sabbagh [32] and Milne and Chalabi [25] proposed different approaches to control its chaotic oscillations.

The (HR) equations are generally presented with the fast time-scale:

$$\begin{aligned}x' &= y - x^3 + 3x^2 - z \\y' &= 1 - 5x^2 - y \\z' &= \varepsilon(4x + K - z)\end{aligned}\tag{5.35}$$

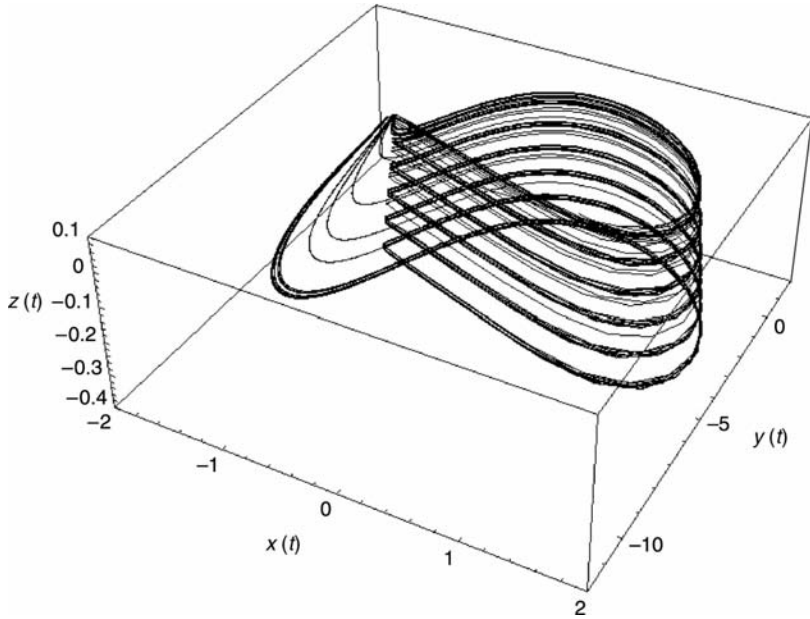
In Equations (5.35), x represents the membrane potential of the neuron, y is the recovery variable, and z is the variable quantifying the mechanism regulating the patterns of discharges [25]. The associated slow system is defined by:

$$\begin{aligned}\varepsilon \dot{x} &= y - x^3 + 3x^2 - z \\ \varepsilon \dot{y} &= 1 - 5x^2 - y \\ \dot{z} &= 4x + K - z\end{aligned}\tag{5.36}$$

where $\varepsilon = 0.004$ and K is a bifurcation parameter. For $K = 3.18$, we observe chaotic bursting oscillations (see the attractor shown in Figure 5.6). The model has only one equilibrium point of the saddle type. This point is defined by solving the equation:

$$x^3 + 2x^2 + 4x + K - 1 = 0\tag{5.37}$$

which has only one real solution.

**FIGURE 5.6**

The chaotic attractor of the HR model obtained by numerical integration for $K = 3.18$ and $\varepsilon = 0.004$, with initial conditions $x(0) = 0.6798$, $y(0) = 0.3825$, and $z(0) = 0.0415$.

The system defined by Equations (5.36) is a singularly perturbed two time-scale model with two fast variables (x and y) and one slow variable (z). The critical manifold M_0 is defined by:

$$\begin{cases} y - x^3 + 3x^2 - z = 0 \\ 1 - 5x^2 - y = 0 \end{cases} \quad (5.38)$$

We deduce the equations defining M_0 :

$$M_0 = \left\{ (x, y, z) \in \mathbb{R}^3 : y = 1 - 5x^2; z = 1 - 2x^2 - x^3 \right\} \quad (5.39)$$

This manifold is the intersection of two cylinders: a parabolic one and a cubic one. It is a curve in the three-dimensional phase space on which the slow dynamics will take place.

Figure 5.7, Figure 5.8, and Figure 5.9, show, respectively, the chaotic temporal evolutions of the HR model variables for $x(t)$, $y(t)$, and $z(t)$ obtained by the numerical integration of Equations (5.36) with initial conditions $x_0 = 0.6798$, $y_0 = 0.3825$, and $z_0 = 0.0415$ [25].

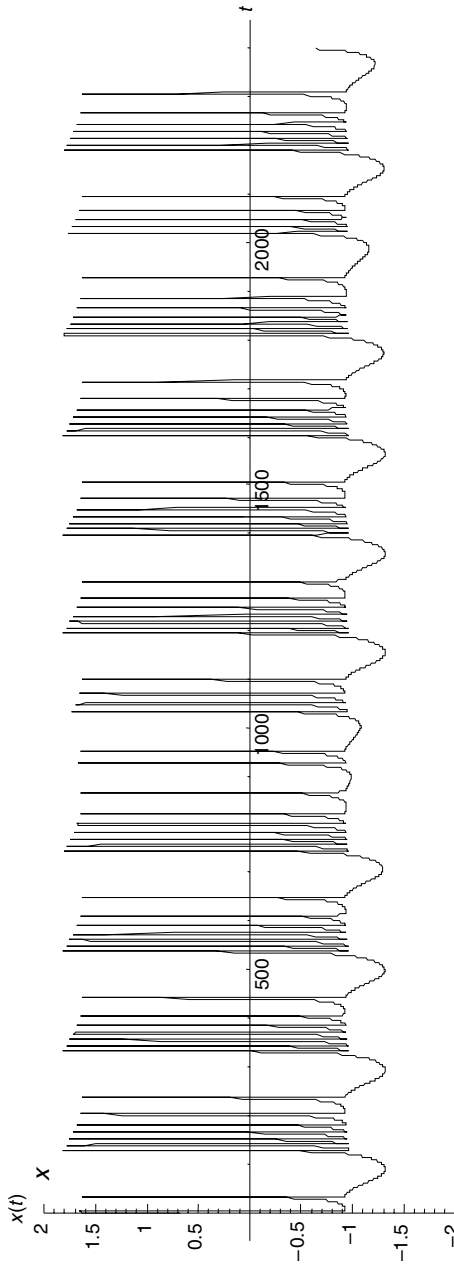


FIGURE 5.7
The chaotic temporal evolutions $x(t)$.

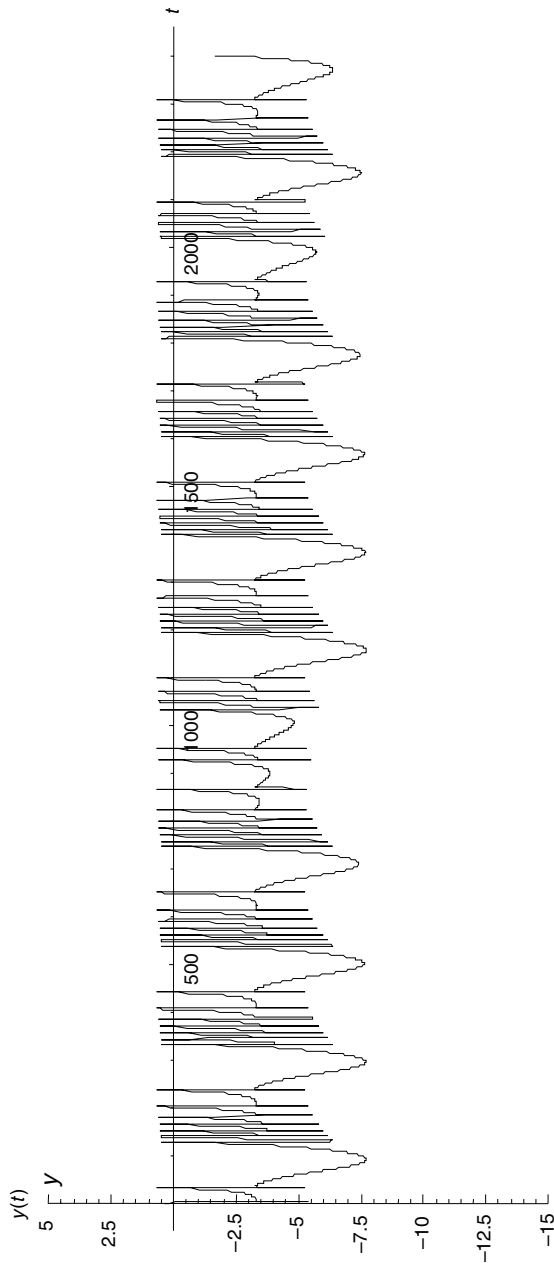


FIGURE 5.8
The chaotic temporal evolutions $y(t)$.

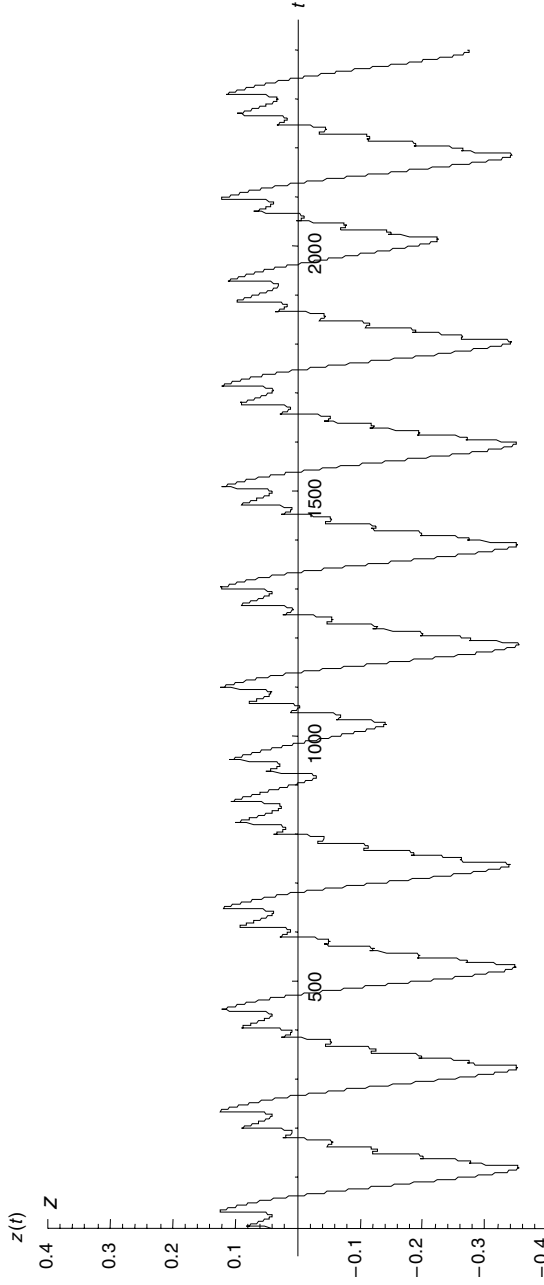


FIGURE 5.9
The chaotic temporal evolutions $z(t)$.

References

1. D.V. Anosov and V.I. Arnol'd, *Dynamical Systems I*, Springer-Verlag, Berlin, 1988.
2. J.P. Barbot, N. Pantalos, S. Monaco, and D. Normand-Cyrot, On the control of singularly perturbed nonlinear systems, in *Proceedings of the IFAC-NOLCOS Symposium*, 453–458, 1992.
3. A. Isidori, *Nonlinear Control Systems: An Introduction*, 2nd ed., Springer-Verlag, Berlin, 1989.
4. P.V. Kokotovic, H.K. Khalil, and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*, Academic Press, London, 1986.
5. R. Marino and P.V. Kokotovic, A geometric approach to nonlinear singularly perturbed control systems, *Automatica*, 24, 31–41, 1988.
6. V.A. Sobolev, Integral manifolds and decomposition of singularly perturbed systems, *Syst. Cont. Lett.*, 5, 169–179, 1984.
7. J. Argémi, Approche qualitative d'un problème de perturbations singulières dans $\mathbb{R}^4$, *Equadiff.* 78, Convegno Int. su equazioni differenziali ed equazioni funzionali, 1978, 333–340.
8. J. Argémi and B. Rossetto, Solutions périodiques discontinues pour l'approximation singulière d'un modèle neurophysiologique dans $\mathbb{R}^4$ — une métaphore dans $\mathbb{R}^3$ avec chaos, *J. Math. Biol.*, 17, 67–92, 1983.
9. V.I. Arnol'd, *Dynamical Systems V*, Springer-Verlag, Berlin, 1994.
10. P. Bergé, Y. Pomeau, and C. Vidal, *L'ordre dans le chaos*, Hermann, Paris, 1984.
11. L.O. Chua, M. Komuro, and T. Matsumoto, The double scroll family, *IEEE Trans. Circ. Syst.*, CAS-33 (11), 1072–1118, 1986.
12. L.O. Chua and G. Lin, Canonical realization of Chua's circuit family, *IEEE Trans. Circ. Syst.*, CAS-37 (7), 885–902, 1990.
13. Y. Fan and A.V. Holden, Bifurcation, burstings, chaos and crises in the Rose–Hindmarsh model for neuronal activity, *Chaos Solitons Fract.*, vol. 3, 439–449, 1993.
14. R. FitzHugh, Impulses and physiological states in theoretical models of nerve membrane, *Biophys. J.*, 1, 445–466, 1961.
15. J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, 5th ed., Springer-Verlag New York, 1997.
16. L. Hindmarsh and R.M. Rose, A model of neuronal bursting using three coupled first order differential equations, *Proc. R. Soc. B*, 221, 87–102, 1984.
17. M.W. Hirsh and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York, 1974.
18. A.L. Hodgkin and A.F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.*, 117, 500–544, 1952.
19. C. KRT Jones, Geometric singular perturbation theory, *Dynamical Systems, Montecatini Terme, Lecture Notes in Mathematics*, vol. 1609, Springer-Verlag, Berlin, 1994, 44–118.
20. H. Kantz and T. Shreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, 1997.

21. T.J. Kaper, An introduction to geometric methods and dynamical systems theory for singular perturbation problems, in *Analyzing Multiscale Phenomena using Singular Perturbation Methods, Proceedings of Symposia in Applied Mathematics*, vol. 56, American Mathematical Society, Baltimore, 1998, 85–131.
22. N. Levinson, Perturbations of discontinuous solutions of nonlinear systems of differential equations, *Acta Math.*, 82, 71–106, 1950.
23. E.N. Lorenz, Deterministic non-periodic flow, *J. Atmos. Sci.*, 20, 130–141, 1963.
24. B.B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, 1985.
25. A.E. Milne and Z.S. Chalabi, Control analysis of the Rose–Hindmarsh model for neural activity, *IMA J. Math. Appl. Med. Biol.*, 18, 53–75, 2001.
26. J.S. Nagumo, S. Arimoto, and S. Yoshizawa, An active pulse transmission line simulating nerve axon, *Proc. IRE*, 50, 2061–2071, 1962.
27. M.I. Rabinovich and H.D.I. Abarbanel, The role of chaos in neural systems, *Neuroscience*, 87, 5–14, 1998.
28. S. Ramdani, Variétés lentes de systèmes dynamiques chaotiques considérés comme lents-rapides — Applications aux Lasers, *Ph.D. Thesis*, 1999.
29. S. Ramdani, B. Rossetto, L.O. Chua, and R. Lozi, Slow manifolds of some chaotic systems with applications to laser systems, *Int. J. Bifurcat. Chaos*, 10 (12), 2729–2744, 2000.
30. B. Rossetto, T. Lenzini, S. Ramdani and G. Suchey, Slow-fast autonomous dynamical systems, *Int. J. Bifurcat. Chaos*, 8 (11), 2135–2145, 1998.
31. D. Ruelle and F. Takens, On the nature of turbulence, *Commun. Math. Phys.*, 20, 167–192, 1971.
32. H. Sabbagh, Control of chaotic solutions of the Hindmarsh–Rose equations, *Chaos Solitons Fract.*, 11, 1213–1218, 2000.
33. X.J. Wang, Genesis of bursting oscillations in the Hindmarsh–Rose model and homoclinicity to a chaotic saddle, *Physica D*, 62, 263–274, 1993.
34. J. Carr, *Application of Center Manifold Theory*, Springer Verlag, New York, 1981.
35. M. Djemai and J.P. Barbot, Singularly perturbed method for the control design of a synchronous motor with its PWM inverter, in *Proceedings of the IEEE Conference on Control Application*, 1995.
36. M. Djemai, J. Hernandez, and J.P. Barbot, Nonlinear control with flux observer for a singularly perturbed induction motor, in *Proceedings of the IEEE Conference on Decision and Control*, 32nd, 1993, 3391–3396.
37. N. Fenichel, Persistence and smoothness of invariant manifolds for flows, *Indiana Univ. Math. J.*, 21, 193–226, 1971.
38. W. Gröbner, *Serie di Lie e Loro Applicazioni*, Poliedro, Cremonese, Roma, 1973.
39. J. Hernandez and J.P. Barbot, Sliding observer-based feedback control for flexible joints manipulators, *Automatica*, 32 (9), 1243–1254, 1996.
40. P.V. Kokotović and P. Sannuti, Singular perturbation method for reducing model order in optimal control design, *IEEE Trans. Automat. Contr.*, 13 (4), 377–384, 1968.
41. D.S. Naidu, Singular perturbation methodology: in control systems, *IEEE Control Engineering Series No. 34*, London U.K., Peter Peregrims LTD, 1988.
42. B. Porter, Singular perturbation methods in the design of full-observers for multivariable linear systems, *Int. J. Contr.*, 26, 589–594, 1977.

43. V. Saksena, J. O'Reilly, and P.V. Kokotovic, Singular perturbations and time-scale methods in control theory: Survey 1976–1983, *Automatica*, 20 (3), 273–293, 1984.
44. M. Spong, K. Khorasani, and P.V. Kokotović, An integral manifold approach to the feedback control of flexible joints robot, *IEEE J. Robot. Automat.*, 3 (4), 291–300, 1987.
45. A.N. Tikhonov, Systems of differential equations containing small parameters multiplying some of derivatives, *Mat. Sub.*, 31, 575–586, 1952.
46. A.B. Vasileva, Asymptotic behaviour of solutions to certain problems involving nonlinear ordinary differential equations containing a small parameter multiplying the highest derivatives, *Russ. Math. Survey*, 18, 13–84, 1963.

Part II

Closed-Loop Design

6

Control of Chaotic and Hyperchaotic Systems

L. Laval

CONTENTS

6.1	Introduction	291
6.2	Chaos Control	294
6.2.1	The OGY Method	294
6.2.2	The Pyragas Method	299
6.2.3	H_∞ -Control of Chaos	302
6.2.4	Adaptive Control of Chaos	304
6.2.5	Sliding Mode Control of Chaos	306
6.2.6	Energy-Based Sliding Mode Control of Chaos	309
6.3	Hyperchaos Control	312
6.3.1	The YLM Method for Hyperchaos Control	312
6.3.2	Enhanced YLM-Method with AAM	315
6.3.2.1	The AAM Mechanism	315
6.4	Conclusions	317
	References	318

6.1 Introduction

Since the seminal work of Ott, Grebogy, and York [43], who demonstrated the ability to stabilize Unstable Periodic Orbits (UPOs) among those embedded in a chaotic attractor, control of chaotic systems has become an extensive field of research.¹ Such an interest is, indeed, motivated by many reasons. For instance, by essence, chaotic systems are highly sensitive to

¹The pioneer work of Hübler [30] is another important starting point in this field.

(tiny) variations of system parameters, initial conditions, and external disturbances. Therefore, in some processes, chaos may lead to harmful or even catastrophic situations² if not subdued. In such cases, the main purpose then is to reduce the chaotic phenomena as much as possible, by means of intentional and suitable control signals applied to the system. As another motivation, in some practical contexts, bringing of chaotic dynamics into a process or exploiting the chaotic nature of a system may efficiently avoid some costly and painful tasks.³ For instance, as pointed out by Ott et al. [43], a chaotic attractor is composed of a dense set of UPOs. Then, a key idea is to associate some of these orbits with different tasks to perform.⁴ Thus, instead of designing, making, and using several devices, the same chaotic system can serve multiple purposes by simply switching, in a controlled manner, among the different orbits of interest. In such a case, the use of chaos then involves stabilizing trajectories of interest,⁵ while preserving, as much as possible, some of the inherent properties of chaotic (possibly hyperchaotic⁶) systems. Another motivation to deal with chaos control is that introduction of particular control laws or modification of the experimental conditions may lead some nonchaotic systems (initially) to perform undesirable chaotic behaviors (i.e., some transitions from order to chaos, sometimes referred to as “chaotification” [13]). Recovery of original properties such as functionality, stability, . . . then implies both design and application of suitable control laws (e.g., see in Gills et al. [25], the “green problem” related to laser systems). Finally, owing to their intrinsic properties (such as sensitive dependence on initial conditions, inherent instability, sensitivity to perturbations and disturbances, etc.) the control of chaos appears to be an interesting challenge to achieve high performance near the stability boundaries.

Such motivations have lead to a huge amount of proposals of chaos control methodologies (e.g., see [3, 13, 22, 36]), and reports on successful applications to experimental processes (e.g., [58]). The current chapter aims at presenting some of these methodologies (which are by no means the only ones valid), hoping to help the readers in understanding the basics regarding this field. In particular, this chapter focuses on two main classes

²For example, if not clearly understood or controlled, the chaotic nature of some chemical reactions may lead to critical instabilities.

³See Ditto and Munakata [19] for some practical applications such as the use of chaotic sensitivity to perturbations to nudge the ISEE-3 spacecraft near a comet with clever burns of fuel.

⁴As illustrative example, some orbits may be associated to performance criteria, in order to characterize some tasks in progress.

⁵In some particular cases, this control purpose may be, nevertheless, reduced to the stabilization of either fixed points or arbitrarily chosen closed orbits (not necessarily related to existing UPOs).

⁶Chaotic systems with more than one positive Lyapunov exponent.

of methods. The first one, essentially due to investigations of physicists and mathematicians, consists of model-independent control methods that aim at directly exploiting the intrinsic properties of chaotic systems to stabilize either (unstable) fixed points or (unstable) periodic orbits. Motivated by suitability to experimental applications, such methods are highly interesting as they usually require limited on-line computation and often make use of experimental procedures to quantify most of the control law parameters for both design and application. Among these methods, this chapter essentially highlights two representative chaos control methodologies: the so-called OGY method ([43, 58]) and Pyragas method [47], as these methods are pioneers in the field and have given rise to many investigations and successful applications (e.g., see [17] for the control of a magneto-elastic ribbon). However, in many cases, this class of methods suffers from the lack of suitable theoretical frameworks to perform rigorous analysis of closed-loop stability [48]. This drawback has, therefore, been a major motivation for introducing the other class under consideration in this chapter, which consists of adapted or extended conventional control techniques coming from the *control theory* framework. However, with regard to the huge number of existing techniques (dedicated to chaos control), the presentation will be restricted to widely known strategies: the H_∞ approach [72], adaptive control [5], and sliding mode control [62, 63], within the context of either dealing with chaotic systems as standard nonlinear ones (i.e., without any direct exploitation of the true nature of chaos) or taking into account some intrinsic properties of chaotic systems (such as the fact that orbits of uncontrolled chaotic systems are confined within a bounded region of the phase space).

This chapter is organized as follows. Section 6.2 deals first with some model-independent control methods. Then, some control strategies based on the control theory framework are presented. Finally, Section 6.3 focuses on some recent methods dedicated to the control of hyperchaotic systems.

PRELIMINARY REMARK

Chaos control methods, presented in the sequel, can be regarded as belonging to the context of “deterministic chaos,” as they essentially consider known (reference) orbits or small neighborhoods of expected system states for either global or local control purposes. However, it is worth mentioning that most of local control approaches [such as the OGY method, the Pyragas method, and (local) H_∞ control] rely implicitly or directly on the properties and results⁷ of ergodic theory (of chaos), which ensures that the system trajectories will always reach expected subsets of the state space where the control has to start acting (effectively). Readers are referred to Appendix A

⁷Such as ergodicity, Poincaré’s recurrence theorem, Birkhoff’s ergodic theorem, etc.

and to the references therein for some complementary informations about such an important and useful probabilistic approach of chaos.

6.2 Chaos Control

Both inherent properties of chaotic processes and uncertainties on system dynamics (including uncertainties on system parameters and neglected dynamics) often complicate experimental applications of chaos control techniques. Therefore, over the past few years, a great deal of attention has been paid to model-independent control methods (i.e., without any refined mathematical description of the chaotic process). In this context, this section starts with the focus on two major techniques: the OGY method⁸ and time-delayed feedback control schemes (i.e., the Pyragas method).

6.2.1 The OGY Method

As previously mentioned, any chaotic system exhibits an extremely dense set of unstable periodic orbits (infinite in number) embedded within the attractor [43]. This fact, combined with the ergodicity⁹ of chaotic orbits, then guarantees that the system trajectory, within the attractor, will always reach (in finite time) close neighborhoods of any point of any UPO. In addition, chaotic systems are highly sensitive to perturbations; therefore, any small disturbance (and, therefore, any control signal) can radically modify their evolution. By considering such properties, Ott, Grebogy, and York proposed [43] a seminal chaos control methodology (referred to as the OGY method) to stabilize some selected UPOs, without drastically altering the inherent chaotic dynamics of the original system. From a technical viewpoint, this methodology essentially consists in constraining the system trajectory to fall on a predetermined stable manifold related to a hyperbolic¹⁰ fixed point of interest, by applying a small time-dependent perturbation on a selected system parameter when the trajectory visits an arbitrarily close neighborhood of the targeted orbit (the so-called OGY region). Thus, to summarize, this methodology involves: (1) the selection of an UPO of interest for stabilization purpose; (2) a suitable framework

⁸The underlying conceptual framework of the OGY method is still an important theoretical basis for investigations, as emphasized by some recent developments about control of high-dimensional and hyperchaotic systems.

⁹A dynamical system is said to be ergodic if the average time spent by a trajectory in any region of the phase space is proportional to the volume of that region (see Birkhoff's ergodic theorem in Appendix A).

¹⁰A saddle point.

for analyzing local stability properties; (3) the design of a control signal according to this entire framework. These steps and the essence of the OGY method can be understood, in a more detailed manner, as follows.

Consider the controlled system defined by the differential equation

$$\dot{x} = f(x, \varepsilon) \quad (6.1)$$

where $x \in \mathbb{R}^n$ is an n -dimensional state vector,¹¹ and f is a vector-valued function which characterizes the system dynamics. $\varepsilon \in \mathbb{R}$ is a system parameter, assumed to be accessible for external adjustment within a small interval around a nominal value ε_0 (i.e., $\varepsilon \in [\varepsilon_0 - |\Delta\varepsilon_{\max}|; \varepsilon_0 + |\Delta\varepsilon_{\max}|]$, where $\Delta\varepsilon_{\max} \in \mathbb{R}$ is the maximum admissible change in the parameter ε to preserve the inherent properties of the original chaotic system). In the following, and without loss of generality, we will consider $\varepsilon_0 \equiv 0$.

In addition, consider a targeted trajectory, as a solution of (6.1) with, for convenience, ε being equal to the nominal value ε_0 . Then, with respect to the (local) stabilization problem under consideration, both capturing of inherent properties of the chaotic attractor structure and setting of a suitable framework for local stability analysis, can be performed by means of the Poincaré section method (see Figure 6.1 and Figure 6.2). Applying this method then leads to the definition of a map (referred to, in the present

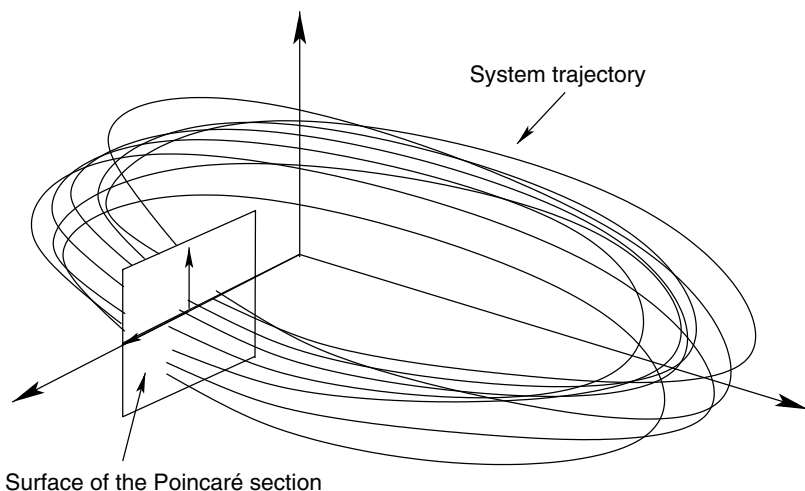
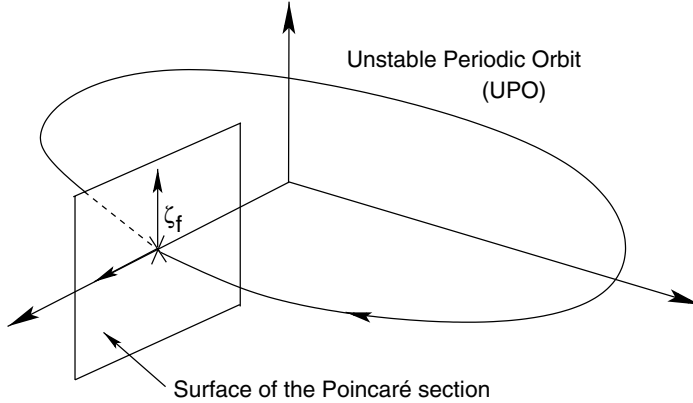


FIGURE 6.1
Chaotic trajectory.

¹¹Ott et al. considered, originally, a three-dimensional system to introduce their control method.

**FIGURE 6.2**

The Poincaré section and an UPO.

context, as a *controlled Poincaré map*) of the form:

$$F : \mathbb{R}^{n-1} \times \mathbb{R} \ni (\zeta, \varepsilon) \mapsto F(\zeta, \varepsilon) \in \mathbb{R}^{n-1}$$

for which ζ represents a point where the orbit intersects the surface of the Poincaré section (when considering one direction for the piercing), F is assumed to be differentiable, and $F(\zeta, \varepsilon)$ stands for the point where the orbit, starting from ζ , first returns to the surface while keeping the input ε constant. Then, iterating the map¹² leads to the following discrete-time dynamical system related to the flow of (6.1),

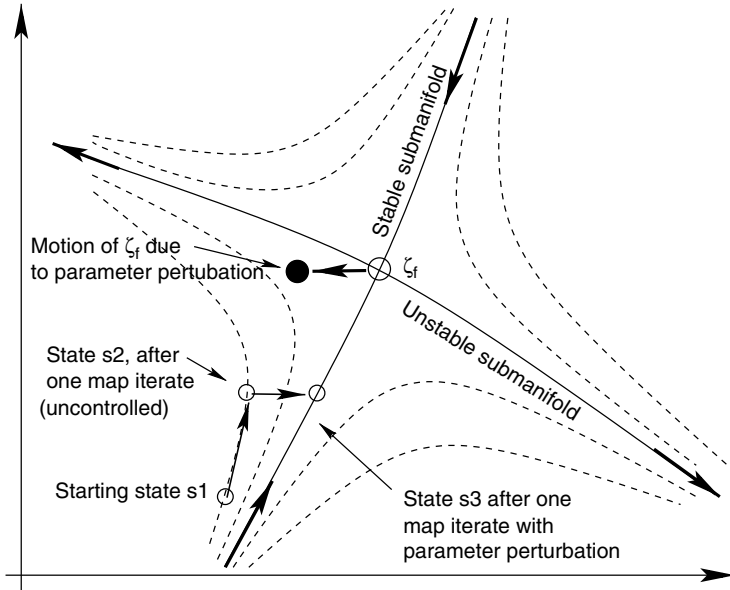
$$\zeta_{k+1} = F(\zeta_k, \varepsilon_k) \quad (6.2)$$

where ζ_k represents the k th intersection of the orbit with the surface (with respect to the piercing direction) and ε_k is the value of ε between ζ_k and ζ_{k+1} corresponding time instants.

Now, without loss of generality, let us assume that the targeted orbit intersects the surface of the Poincaré section in only one fixed point¹³ assumed to be hyperbolic, and let us denote by ζ_f this fixed point (assumed to exist for $\varepsilon = \varepsilon_0$; i.e., $F(\zeta_f, \varepsilon_0) = \zeta_f$) (see Figure 6.2 and 6.3). Moreover, without loss of generality, let us assume that $\zeta_f = 0$. Then, due to sensitivity of chaotic systems to small perturbations, any change in the parameter ε induces a shift of the fixed point coordinates correspondingly. Thus, with

¹²Such an operation leads any continuous-time periodic orbit to appear as a discrete-time orbit cycling through a finite set of points lying on the surface of the section (where the number of crossing points, with respect to the piercing direction, depends on the periodicity of the orbit).

¹³Hence, the targeted orbit under consideration here is of period one.

**FIGURE 6.3**

Schematic explanation of OGY method: curved trajectories follow the stable (sub)manifold toward the periodic point and the unstable (sub)manifold away from this point. Without any parameter perturbation, the state s_1 evolves to the state s_2 . The effect of changing the parameter ε is depicted as shifting the states near ζ_f along the solid black arrows. The combination of the unperturbed trajectory and the effect of the perturbation induces a shift of the fixed point and its (sub)manifolds, so that the next iterate, s_3 , falls on the stable (sub)manifold. Once on the stable (sub)manifold, the trajectory naturally tends toward the desired periodic orbit.

some mild assumptions on both ε and ζ_f variations, the OGY method suggests to linearize the Poincaré map about the desired fixed point ζ_f and the nominal parameter ε_0 , so as to obtain a local linearized model of the form,¹⁴

$$\Delta \zeta_{k+1} = A \Delta \zeta_k + B \Delta \varepsilon_k \quad (6.3)$$

where A is the Jacobian matrix of $F(\bullet, \varepsilon_0)$ evaluated at ζ_f , $B = \partial F / \partial \varepsilon(\zeta_f, \varepsilon_0)$ is the derivative of F with respect to the parameter ε , $\Delta \zeta_k = \zeta_k - \zeta_f$, and $\Delta \varepsilon_k = \varepsilon_k - \varepsilon_0$.

Then, as previously mentioned, assuming the fixed point is a saddle, there exist two eigenvalues of the surface at ζ_f , λ_s , and λ_u , which satisfy $|\lambda_s| < 1 < |\lambda_u|$ (where the subscripts s and u stand, respectively, for stable and unstable). Accordingly, there exist two associated right eigenvectors, w_u and w_s , defined as $w_u^T A = \lambda_u w_u^T$ and $w_s^T A = \lambda_s w_s^T$, and such

¹⁴Recall that, for convenience but without loss of generality, ζ_f and δ_0 are assumed to be such that $\zeta_f = 0$ and $\delta_0 = 0$.

that $v_u^T w_u = 1$, $v_s^T w_s = 1$, $v_u^T w_s = 0$, $v_s^T w_u = 0$, where v_u and v_s are two left eigenvectors satisfying $Av_u = \lambda_u v_u$ and $Av_s = \lambda_s v_s$.

Finally, assuming ζ_k be located within a neighborhood of the desired fixed point, the control problem then consists in selecting ε_k so that ζ_{k+1} be put onto the stable manifold (i.e., to select ε_k so that $w_u^T \Delta \zeta_{k+1} = 0$). For this purpose, dotting (6.3) with w_u leads, finally, to define a control law of the form¹⁵ [43, 70],

$$\Delta \varepsilon_k = -\lambda_u \frac{w_u^T \Delta \zeta_k}{w_u^T B} \quad (6.4)$$

which can be used to design the (small) control signal, provided the magnitude of the right-hand side is in $[\varepsilon_0 - |\Delta \varepsilon_{\max}|; \varepsilon_0 + |\Delta \varepsilon_{\max}|]$.

Otherwise ε_k has to be set to zero (i.e., $\varepsilon_k = \varepsilon_0$), because the control may not be able to bring the orbit to the fixed point.

Finally, let us express some technical and concluding remarks:

- When ζ_{k+1} falls on the stable submanifold (related to the fixed point of interest), the parameter ε can be set to its nominal value, because subsequently the orbit will approach the desired fixed point.
- A major problem occurs for the incomplete measurement of the system state. However, this problem can be overcome by replacing the initial state vector x by the so-called delay coordinate vector¹⁶ (see [44, 59]) of the form:

$$X(t) = [y(t), y(t - \tau), \dots, y(t - (m - 1)\tau)]^T \in \mathbb{R}^m$$

where y is a (noise-free) system output available for measurement, m corresponds to the embedding dimension, and $\tau > 0$ is a (selected) time delay (i.e., an embedding time). One can refer to the literature [53, 60] for further details on the selection of parameters m and τ .

- The original OGY methodology considers saddle type fixed points that have both stable and unstable submanifolds. However, as pointed out by Yu et al. [69], the construction of stable and unstable (sub)manifolds for high dimensional chaotic systems is a technical challenge [40]. Thus, the OGY method appears to be mainly suitable for the control of lower dimensional chaotic systems.

¹⁵Recall that $\Delta \varepsilon_k = \varepsilon_k - \varepsilon_0$, with ε_0 assumed to be zero in the present case.

¹⁶Indeed, Takens [59] pointed out that, for sufficiently large embedding dimensions and time delay, there exists a smooth and invertible mapping between the actual flow and the reconstructed flow (by means of time delay coordinates).

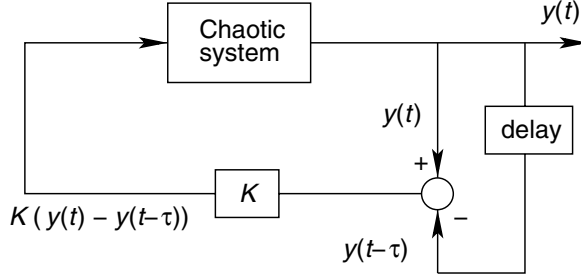
- Ergodicity of chaotic orbits guarantees the reach, in finite time, of a neighborhood of the targeted orbit (see both Poincaré's recurrence theorem and Birkhoff's ergodic theorem in Appendix A). However, before the reaching of this neighborhood, the system evolves in a chaotic mannered (i.e., similar to an uncontrolled chaotic system). Thus, the waiting time can be quite long as it depends on both the (arbitrary) chosen initial conditions and the selected UPO (for stabilization purpose). However, to overcome this problem, various targeting methods have been proposed for reducing the waiting time (see [32] and references therein).
- The OGY method is based on the linear approximation of the Poincaré map, which does not capture the nonlinear dynamics of the system. Therefore, in some cases (e.g., highly nonlinear systems), the control may not be able to bring the orbit to the fixed point, leading the system trajectory to leave the OGY region for wandering chaotically (as if there were no control) until the next reach of the neighborhood of interest.¹⁷

The OGY method has given rise to many extended or modified local control schemes which cannot be fully listed here (for detailed surveys, readers can refer to [13, 18, 28] and references therein). For instance, Yu et al. [69, 70] extended the OGY method to deal with high-dimensional systems by removing the reliance of the control on eigenvalues and eigenvectors of the system Jacobians. Aston and Bird [4] studied the extension of the *immediate basin of attraction* (i.e., the OGY region) by considering further regions. Romeiras et al. [49] proposed an extension of the method so that the control problem can be addressed by means of classical state feedback control design approaches, allowing for application to high-dimensional systems. Finally, it is worth mentioning a seminal modification, referred to as occasional proportional feedback ([31, 45], see also [66] and references therein for detailed information), which has lead to several successful applications (e.g., control of a chaotic state laser system [51], synchronization of diode resonators [42], and stabilization of chaotic fluctuations in the frequency emission from a tunable-diode laser [15]).

6.2.2 The Pyragas Method

Pyragas [47] proposed an alternative model-independent control method to the OGY method, which gained widespread acceptance as it does not require any real-time computation and has been applied to many

¹⁷This reach being guaranteed by the ergodicity property of the chaotic system (see Appendix A).

**FIGURE 6.4**

Schematic representation of the Pyragas control scheme.

experimental processes including electronic, lasers, chemical systems, hydrodynamical systems, and cardiac systems (e.g., see some references in [3, 22]). This method, commonly referred to as the time-delayed autosynchronization method (TDAS), the time-delayed feedback control method (TDFC), or the Pyragas method, simply requires the accessibility to a scalar variable for measurement and, at least, one system input for external control, so as to perform a control scheme as depicted in Figure 6.4.

From a technical viewpoint, this control scheme can be understood by considering a mathematical description of the form,

$$\dot{z}(t) = Q(y(t), z(t)) \quad (6.5)$$

$$\dot{y}(t) = P(y(t), z(t)) + u(t) \quad (6.6)$$

where $y(t)$ is the output scalar variable (accessible for measurement), $z(t)$ represents the remaining hidden variables of the dynamical system,¹⁸ $u(t)$ is the input signal, and P and Q are two vector-valued functions of appropriate dimensions.

Then, the use of the standard method of time-delay coordinates affords the possibility to extract, from measured variable y , various periodic signals of the form $y(t) = y_i(t)$, $y_i(t + T_i) = y_i(t), \dots$ where T_i represents the “period” of the i th UPO. To perform the stabilization of a selected UPO, the Pyragas method then consists in considering a simple external feedback control action $u(t)$ of the form,

$$u(t) = K[y(t - \tau) - y(t)] \quad (6.7)$$

where K is a negative feedback gain and τ represents a time delay.

For a noise free, measured signal, stabilization of the i th UPO is then achieved when τ equals the period T_i (remark: in such a case, the control signal also vanishes).

¹⁸The whole state vector $x(t)$ is $x(t) = [z(t) y(t)]^T$.

Finally, let us express some technical and concluding remarks:

- The Pyragas method has given rise to many investigations. For instance, Socolar et al. [56] proposed an extended version, referred to as ETDAS, which affords the stabilization of orbits that the basic TDAS cannot [1, 56]. This extension consists in replacing Equation (6.7) with a control law of the form,

$$u(t) = K \left[y(t) - (1 - R) \sum_{k=1}^{\infty} R^{k-1} y(t - k\tau) \right], \quad k \in \mathbb{N} - \{0\} \quad (6.8)$$

for which, $R \in [0, 1)$ is a suitable real parameter ($R = 0$ corresponds to the basic TDAS). However, for continuous-time systems, both numerical results and experiments have shown that for ETDAS to be successful, the feedback gain K must lie within a finite, and often narrow range [10].

- As pointed out by Arecchi et al. [3], the use of control laws (6.7) or (6.8) transforms the original system equations into delayed differential ones. As a delayed system is richer than an instantaneous one, care should be taken in stabilizing a true UPO of the original unperturbed system rather than a spurious UPO introduced by the delay.
- Within the Pyragas-type control context, rigorous analysis of the closed loop behavior is not trivial (e.g., see [10] for a stability analysis of ETDAS control scheme). Therefore, until recently, only numerical and experimental results concerning performance and limitations of the Pyragas method were proposed. However, Basso et al. [7, 8] introduced a bridge between the Pyragas approach and the classical feedback control theory by formulating the periodic orbit stabilization as an input–output L_2 stability problem of a linear periodic feedback system. Such a framework then makes possible the use of some criteria in the L_2 -setting (circle criterion, Willems criterion, etc.) to determine the stability bounds of the periodic orbit. Following the same purpose, Ushio [61] established, for a class of discrete-time systems, a necessary condition for stabilizability with a Pyragas controller (6.7). This result was extended for more general and continuous-time cases by Just et al. [35] and Nakajima [41]. These works also pointed out a limitation of the basic Pyragas method as well as its various modifications. This limitation is that any UPO with odd number of Floquet multipliers¹⁹

¹⁹Given a periodic orbit Γ , a Poincaré map P transverse to Γ , and a fixed point ζ_f of Γ lying on the surface associated to P , the *Floquet multipliers* are the eigenvalues of the linearization of P evaluated at ζ_f . Therefore, Floquet multipliers provide a method to analyze the stability of the periodic orbit.

whose modulus is greater than unity can never be stabilized by the delayed feedback control method. However, recently, this drawback was overcome by introducing, into the feedback loop, an unstable degree of freedom to artificially enlarge the set of these multipliers to an even number [48].

- Finally, as pointed out in [22], control law (6.7) is sensitive to the selection of the time delay τ . Then, when the mandatory choice $\tau = T$ cannot be efficiently considered, an alternative empirical approach is to simulate the uncontrolled system with initial condition $x(0)$ until the current state $x(t)$ reaches a close neighborhood of a desired state $x(s)$. Then, Equation (6.7) can be exploited by considering both $\tau = t - s$ as a reasonable estimate of the period, and vector $x(t)$ as initial condition to start the control.

Both the OGY and the Pyragas methods are model-independent techniques that aim at stabilizing the periodic orbits by exploiting intrinsic properties of chaos. However, taming chaos can also be considered from the particular viewpoint of control theory, by means of strategies based on nominal, possibly uncertain, modeling of chaotic systems. Such a consideration has led to a huge amount of proposals of control schemes, representatives of which are presented in the following sections. This presentation is nevertheless restricted here to some methods capable of dealing with one or more conditions that complicate experimental applications, namely: uncertainties on system parameters, neglected dynamics, noise affecting the measured signals, etc.

6.2.3 H_∞ -Control of Chaos

As established by Romeiras et al. [49], the OGY method can be extended so that the stabilization problem can be addressed by means of conventional state feedback control. Then, with regard to this conceptual framework, Jonckheere et al. [34] studied LQ control of chaos. Moreover, as a natural extension of this work, Bhajekar et al. [9] investigated the design of feedback controllers within the context of the H_∞ approach (as a well-suited framework to deal with robust control problems [74]). This has led to an H_∞ (linear) control scheme whose strategy is presented as follows.

With regard to the formal framework of the OGY method, consider a chaotic system defined by a recurrence equation²⁰ of the form $x(k+1) = F(x(k), \varepsilon)$, ($k \in \mathbb{N}$), where $x \in \mathbb{R}^n$ is the state vector and $\varepsilon \in \mathbb{R}$ is a system parameter assumed to be accessible for external adjustment within a small interval around a nominal value ε_0 . Moreover, let x_f be the desired fixed point to which the system is intended to be driven. In addition, assume

²⁰Related, for instance, to a Poincaré map.

that a nominal, local, linear model of the actual system can be obtained from linearization about the desired fixed point x_f (lying on a Poincaré section surface) and nominal parameter ε_0 . Then, by considering an error signal $z(k)$ as the weighted sum of the state error and the control effort, and an additive perturbation $a(k)$ to characterize the neglected nonlinear dynamics, the controller design problem can be based on a (generalized) plant description of the form [9],

$$\begin{aligned}\tilde{x}(k+1) &= A \tilde{x}(k) + B_1 a(k) + B_2 \tilde{\varepsilon}(k) \\ z(k) &= C_1 \tilde{x}(k) + D_{11} a(k) + D_{12} \tilde{\varepsilon}(k) \\ y(k) &= C_2 \tilde{x}(k) + D_{21} a(k) + D_{22} \tilde{\varepsilon}(k)\end{aligned}\tag{6.9}$$

where $\tilde{x}(k) = x(k) - x_f$, $\tilde{\varepsilon}(k) = \varepsilon(k) - \varepsilon_0$, $x(k) \in \mathbb{R}^n$ denotes the state of the system (which may not be fully measurable); $y(k)$ is the measured output; $a(k)$ is the exogenous input; $z(k)$ is a performance measure (i.e., the controlled error); $\varepsilon(k)$ is the control input; and $A, B_1, B_2, C_1, C_2, D_{11}, D_{12}, D_{21}$, and D_{22} are constant matrices of appropriate dimensions.

As considered in [9], assume that $B_1 = 1$, $C_2 = 1$, $D_{11} = 0$, $D_{12} = I$, $D_{21} = D_{22} = 0$. Moreover, in this context, consider the following usual assumptions:

- A1. A is assumed to be nonsingular (and therefore invertible)
- A2. (A, B_2) is stabilizable and (C_2, A) is detectable

Then, by considering the closed-loop transfer function T_{za} which maps a to z , the robust H_∞ feedback control problem can be expressed as finding a (static) linear controller K_∞ such that:

$$\tilde{\varepsilon}(k) = K_\infty \tilde{x}(k)\tag{6.10}$$

$$\|T_{za}\|_\infty \leq \gamma \quad \text{with} \quad 0 < \gamma \leq \gamma_{\text{optimal}}\tag{6.11}$$

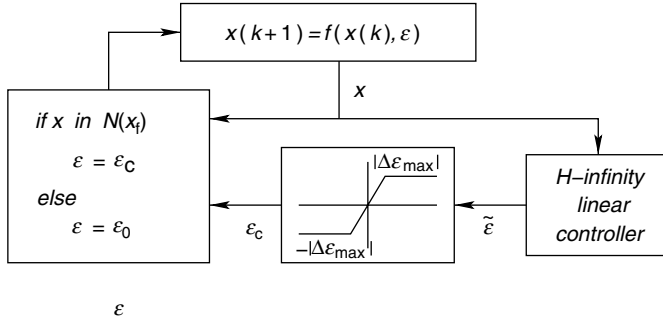
Then, according to [6], the suboptimal H_∞ controller which ensures that $\|T_{za}\|_\infty < \gamma_{\text{optimal}}$ is given by,

$$\tilde{\varepsilon}(k) = -B^T P \Lambda^{-1} A \tilde{x}(k)\tag{6.12}$$

where $\Lambda = I + (BB^T - \gamma^{-2}I)P$, and P is the positive definite solution of the Generalized Algebraic Riccati Equation,

$$C_1^T C_1 + A^T P \Lambda^{-1} P - P = 0\tag{6.13}$$

with the existence condition $\gamma^2 I - P > 0$.

**FIGURE 6.5**

The (linear) H_∞ control scheme, including a nonlinear saturation element to ensure admissible variations of ε (so that $\varepsilon \in [\varepsilon_0 - |\Delta\varepsilon_{\max}|; \varepsilon_0 + |\Delta\varepsilon_{\max}|]$, $\Delta\varepsilon_{\max} \in \mathbb{R}$).

Finally, by considering that the measured output signal $x(k)$ lies within an arbitrarily close neighborhood²¹ $N(x_f)$ of the fixed point x_f , the H_∞ control strategy (as proposed in [9]) can then be summarized as shown in Figure 6.5.

This H_∞ control scheme can be viewed as a sort of bridge between control theory and OGY-type methods as it considers, in part, the same conceptual framework.²² However, note that, according to the control algorithm, the control action takes place when, with respect to the ergodicity property, the free system trajectory reaches a close neighborhood of the fixed point to stabilize, assuming that this neighborhood becomes invariant by means of the H_∞ control action. Thus, when compared with conventional linear H_∞ control, the current control scheme is of an unusual form as it involves a (robust) domain of attraction restricted here to local considerations.

6.2.4 Adaptive Control of Chaos

In 1989, the pioneering work of Hübler [30] demonstrated some capabilities of an adaptive control scheme for dealing with chaotic systems. Since then, as a conventional and well-developed approach for controlling either certain or uncertain dynamical systems, adaptive and adaptive-like techniques have received a great deal of attention for the control of chaos [13, 23]. For instance, a two steps adaptive-like control scheme was proposed by Arecchi et al. [2, 11] which consists in: (1) extracting both unperturbed features of system dynamics and periods of the

²¹It is worth mentioning again that the reach of such a neighborhood relies on the ergodic property of chaotic systems (see Appendix A).

²²For instance, by considering the application of a control signal on a selected system parameter when the trajectory visits an arbitrarily close neighborhood of the targeted orbit.

UPOs and (2) designing and applying adaptive perturbations to stabilize a selected UPO. Zhang et al. [71] investigated an adaptive control based on the Gaussian radial basis function networks²³ [52]. Closer approaches to classical adaptive control theory were also proposed [16, 20, 23, 26, 27] to deal with unmodeled dynamics, unknown parameters, and disturbances. In this section, we consider an adaptive strategy as proposed in [16], which summarizes the essence of a basic Model Reference Adaptive Control (MRAC) scheme for chaotic systems. More precisely, consider the following two systems, where the first one represents the system to control and the second one is used as a reference model to characterize the expected dynamical behavior:

$$\dot{x}(t) = f(x, t) + Bu(t), \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \quad B \in \mathbb{R}^{n \times m} \quad (6.14)$$

$$\dot{y}(t) = g(y, t), \quad y \in \mathbb{R}^n \quad (6.15)$$

The control problem is then to choose an appropriate (adaptive) control law $u(t)$ such that, $\lim_{t \rightarrow \infty} \|x(t) - y(t)\| = 0$, where $\|\bullet\|$ is the Euclidean norm.

With this aim, consider the (adaptive) control strategy proposed by Di Bernardo [16]. This strategy consists in first rewriting the control problem above in an appropriate form, so that standard control design approaches could be applied. This leads to the expression of the error dynamics as,

$$\dot{e}(t) = \dot{x}(t) - \dot{y}(t) = f(x, t) - g(y, t) + Bu(t) \quad (6.16)$$

Now, assume that there exists an appropriate orthogonal projection operator $\Pi : \mathbb{R}^n \rightarrow \text{Im}(B)$ [16], so that relation (6.16) can be rewritten as

$$\dot{e}(t) = Le(t) + B[h(x, t) - l(y, t) + u(t)] \quad (6.17)$$

where $Le(t)$ is the projection of $f(x, t) - g(y, t)$ on the complementary space of $\text{Im}(B)$, which is assumed to be linear, and $h(x, t), l(y, t)$ are the projection on $\text{Im}(B)$ of $f(x, t)$ and $g(y, t)$, respectively.

By noting that relation (6.17) is of an appropriate form to deal with the design of a state feedback controllers, the problem can be expressed as finding a gain matrix K of appropriate dimension and such that $\tilde{L} = L - BK$ is Hurwitz. Then, exploiting the fact that the reference model trajectory is restricted to be a chaotic orbit, a limit cycle, or an equilibrium point (hence, its evolution is bounded), one can consider a control law of the form [16],

$$u(t) = -Ke(t) - k(t)[1 + \phi(x)] \|B^T Pe\|^{-1} B^T Pe \quad (6.18)$$

²³A neural network approach.

where P is the solution of the Lyapunov equation $P\tilde{L} + \tilde{L}^T P + I = 0$, $\phi(x)$ is a positive definite, continuous function which defines an upper-bound of the nonlinearity of the system to control, and $k(t)$ is adaptively estimated according to the law,

$$\dot{k}(t) = [1 + \phi(x)] \|B^T P e\| \quad (6.19)$$

which guarantees that the error $e(t)$ asymptotically converges towards zero (see the proof in Di Bernardo [16]).

REMARK 1

Control law (6.18) consists of two different contributions: a linear feedback term and a discontinuous action whose magnitude is adaptively estimated. As pointed out earlier [39], the discontinuous action may induce an unfavorable chattering phenomenon. However, this problem vanishes if, in Equation (6.17), the linear matrix L is Hurwitz. Indeed, in such a case, the linear feedback term can be omitted, leading to a pure adaptive controller of the form,

$$\begin{cases} \dot{k}(t) = \|B^T P e\| \\ u(t) = -k(t) \|B^T P e\|^{-1} B^T P e \end{cases} \quad (6.20)$$

6.2.5 Sliding Mode Control of Chaos

First, recall that conventional sliding mode control [62] requires at least: (1) a switching manifold that prescribed the desired dynamics and (2) a discontinuous control law such that the system trajectory first reaches the manifold and then stays on it forever. Within this context, as a preliminary approach, Vincent and Yu [65] investigated the use of a bang-bang controller to stabilize one of the unstable equilibrium points. This kind of control was also studied by Galiás and Orgozalek [24] as a modification of the OGY method in case of assuming only two values for the control parameter. Later, Yu [68] introduced a sliding mode control strategy based on the switching-in of a two-value Lorenz system parameter. More recently, Yu et al. [70] emphasized the interest for the variable structure control approach, by proposing an extension of the OGY-type control method based on invariant manifold theory. Finally, it is worth mentioning the work of Yau et al. [64] who investigated sliding mode control for a class of chaotic systems with uncertainties. As main features, their control scheme: (1) guarantees asymptotic tracking of stable or unstable periodic orbits; (2) avoids the so-called “chattering effect”; and (3) does not require either the explicit use of a Poincaré map or the linearization about some specific point. This section deals with this last work, to apprehend some capabilities of sliding mode control methods to deal with chaotic systems (e.g., [33]).

For this purpose, consider the class of chaotic systems defined by

$$\begin{aligned}\dot{x}_i &= x_{i+1}, \quad 1 \leq i \leq n-1 \\ \dot{x}_n &= f(x, t) + \Delta f(x) + u\end{aligned}\tag{6.21}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}$ is the control input, f is a given non-linear C^2 -function, and $\Delta f(x)$ is a twice (continuously) differentiable uncertain term which represents the unmodeled dynamics and/or structural variations of the system.

REMARK 2

The class under consideration consists of systems (with or without uncertainties) which can be expressed in a well-defined triangular form which ensures the so-called matching condition.²⁴ For instance, such a class includes the Duffing–Holmes damped spring system, Van der Pol oscillator, etc.

Moreover, assume that:

- A1. $f(x, t)$ and $\Delta f(x)$ satisfy all the necessary conditions such that system (6.21) has a unique solution in the time interval $[t_0; +\infty)$, $t_0 > 0$, for any given initial condition $x_0 = x(t_0)$
- A2. System (6.21), with $u = 0$, evolves in a chaotic motion

In this context, the control problem can be formulated as finding a sliding mode control law $u(t)$ such that: $\lim_{t \rightarrow \infty} \|x(t) - \tilde{x}(t)\| \rightarrow 0$, where $\|\bullet\|$ is the Euclidean norm and $\tilde{x}(t)$ denotes a targeted orbit. For this purpose, consider $\tilde{x}(t)$ as a solution of the unperturbed and uncontrolled system,

$$\begin{aligned}\dot{\tilde{x}}_i &= \tilde{x}_{i+1}, \quad 1 \leq i \leq n-1 \\ \dot{\tilde{x}}_n &= f(\tilde{x}, t)\end{aligned}\tag{6.22}$$

and let the tracking error be $e_j = x_j - \tilde{x}_j$, $j = 1, 2, \dots, n$, and $g(e, \tilde{x}, t) = f(e + \tilde{x}, t) - f(\tilde{x}, t)$.

Subtracting (6.22) from (6.21) leads to the characterization of error dynamics as,

$$\begin{aligned}\dot{e}_i &= e_{i+1}, \quad 1 \leq i \leq n-1 \\ \dot{e}_n &= g(e, \tilde{x}, t) + \Delta f(e + \tilde{x}) + u\end{aligned}\tag{6.23}$$

where $g(e, \tilde{x}, t)$ and $\Delta f(e + \tilde{x})$ are assumed to be bounded.

²⁴So that the uncertainties can be compensated directly by means of the control law $u(t)$.

Then, assuming that u is differentiable and that first derivatives of both $g(e, \tilde{x}, t)$ and $\Delta f(e + \tilde{x})$ are bounded, relation (6.23) can be rewritten, by the use of extended systems concept [21], as a standard system of the form:

$$\begin{aligned} \dot{e}_i &= e_{i+1}, \quad 1 \leq i \leq n-1 \\ \dot{e}_n &= g(e, \tilde{x}, t) + \Delta f(e + \tilde{x}) + u \equiv e_{n+1} \\ \dot{e}_{n+1} &= \frac{d}{dt} [g(e, \tilde{x}, t) + \Delta f(e + \tilde{x})] + \dot{u} \end{aligned} \quad (6.24)$$

By noting that system (6.24) is of a (generalized) canonical controllable form [21] without any internal dynamics, and by considering the work of Chen and Lin [14], the (extended) sliding surface can then be defined by,

$$s = e_{n+1} - e_{0(n+1)} + \int_0^t \sum_{j=1}^{n+1} c_j e_j dt = 0 \quad (6.25)$$

where $e_{0(n+1)}$ denotes the initial state of e_{n+1} .

Now, assuming that $s = 0$ and initial condition $e_{n+1}(0) = e_{0(n+1)}$, the sliding mode dynamics can then be described by the following set of equations,

$$\begin{aligned} \dot{e}_i &= e_{i+1}, \quad 1 \leq i \leq n \\ \dot{e}_{n+1} &= - \sum_{j=1}^{n+1} c_j e_j \end{aligned} \quad (6.26)$$

for which the design parameters c_j can be determined so that the characteristic polynomial

$$P(e) = \dot{e}_{n+1} + \sum_{j=1}^{n+1} c_j e_j \quad (6.27)$$

is Hurwitz.

Then, according to [55], the reaching law can be chosen as,

$$\dot{s} = -w \operatorname{sign}(s) \quad (6.28)$$

where $\operatorname{sign}(\cdot)$ denotes the sign function and the switching gain $w > 0$ has to be determined such that the sliding condition is satisfied and sliding mode motion will occur.

Finally, gathering relations (6.25), (6.26), and (6.28) leads to a control law $\dot{u}$ of the form,²⁵

$$\dot{u} = -\frac{d}{dt} [g(e, \tilde{x}, t) + \Delta f(e + \tilde{x})] - w \operatorname{sign}(s) - \sum_{j=1}^{n+1} c_j e_j \quad (6.29)$$

However, recall that the uncertainties $\Delta f(e + \tilde{x})$ are assumed to be unknown (but upper bounded). Therefore (6.29) cannot be used directly for implementation. Instead, assuming the control to be robust enough to tackle some dynamics which are not directly compensated, (6.29) can be replaced by,

$$\dot{u} = -\frac{d}{dt} (g(e, \tilde{x}, t)) - w \operatorname{sign}(s) - \sum_{j=1}^{n+1} c_j e_j$$

REMARK 3

Due to the control effect, on the sliding manifold, the error e_n tends towards zero and, therefore, both $\dot{e}_n$ and $g(e, \tilde{x}, t)$ also converge to zero. Thus, the control law $u(t)$ is always bounded.

As shown by realistic simulation results [64], the sliding mode control appears to be suitable for chaos control purposes. In particular, robustness property of sliding mode controllers affords dealing with uncertain and/or perturbed chaotic systems, provided an upper bound of these uncertainties and/or disturbances is known. However, note that standard (i.e., first order) sliding mode control has some drawbacks such as the well-known chattering effect, which may lead to enforce the instability of the chaotic system (under control). Thus care should be taken in designing a sliding mode control law for such systems.

6.2.6 Energy-Based Sliding Mode Control of Chaos

As mentioned by Prigogine and Stengers [46], the more complex a system is, the more numerous are the perturbations, disturbances, or fluctuations that threaten its stability. As the system becomes more vulnerable to these disturbances, its energy requirement increases as it tries to maintain its structural properties. In contrast, any chaotic system can be considered as having its own limited energy source. Therefore, by controlling only one or several states, the system may be able to stabilize itself by using its own energy source. Such notions are interesting enough to motivate the

²⁵Recall that the first derivatives of both $g(e, \tilde{x}, t)$ and $\Delta f(e + \tilde{x})$ are assumed to be bounded.

consideration of the system energy for both designing control laws and taming chaos. Following this purpose, an energy-based control strategy for chaotic (possibly hyperchaotic) systems was proposed [37, 38]. This strategy mainly consists in a Variable Structure Control (VSC) approach which considers explicitly the system energy for both controller design and system stabilization. More precisely, with this strategy, the control objective is to regulate the system energy with respect to a shaped, nominal representation, implicitly related to system trajectories, robustness, and performance. From a more technical viewpoint, this control approach can be presented as follows.

Let us consider the class of n -dimensional, nonlinear, autonomous systems of the form,

$$\begin{aligned}\dot{x} &= f(x) + u \\ y &= h(x)\end{aligned}\tag{6.30}$$

where $x \in \mathbb{R}^n$ is the state vector, f is a vector-valued function, $u \in \mathbb{R}^n$ is the vector of control inputs, and $y \in \mathbb{R}^m$ is the output vector. Then, let us introduce some technical assumptions:

- A1. State vector $x \in \mathbb{R}^n$ can be partitioned as $x = [x_1^T \ x_2^T]^T$ with $x_1 \in \mathbb{R}^{n-2}$, $x_2 = [x_{21} \ x_{22}]^T \in \mathbb{R}^2$, and $f(x)$ can be written as $f(x) = [f_1^T(x_1, x_2) \ f_2^T(x_1, x_2)]^T$ with $f_1(x_1, x_2) \in C^1(\mathbb{R}^{n-2})$ and $f_2(x_1, x_2) \in C^1(\mathbb{R}^2)$

REMARK 4

The dimension of x_2 is restricted here to a two-dimensional vector for the convenience of bifurcation analysis.

- A2. The system is at least locally observable and controllable.

In addition, let us consider that the system energy can be represented by a Lyapunov function V which can be divided into two parts V_1 and V_2 (i.e., $V = V_1 + V_2$) related to scalar positive functions $V_T(x_1)$ and $V_{IS}(x_2)$, respectively. Moreover, assume that the positive functions $V_T(x_1)$ and $V_{IS}(x_2)$ have continuous first derivatives which can be expressed as:

$$\dot{V}_T = \dot{x}_1^T \Psi_1 x_1\tag{6.31}$$

$$\dot{V}_{IS} = \dot{x}_2^T \Psi_2 x_2\tag{6.32}$$

where $\Psi_1 \in \mathbb{R}^{(n-2) \times (n-2)}$ and $\Psi_2 \in \mathbb{R}^{2 \times 2}$ are diagonal matrices with strictly positive real values.

Then, with respect to this context, the energy-based sliding mode control can be defined by the following result (see Laval and M'Sirdi [38] for the proof).

LEMMA 1

Consider the following control structure,

$$u = \begin{bmatrix} u_T^T & u_{IS}^T \end{bmatrix}^T \quad (6.33)$$

with,

$$u_T = -\Gamma_1 \text{sign}(V_T) x_1 - \hat{F}_1(x_1, x_2) \quad (6.34)$$

$$u_{IS} = -\Gamma_2 \text{sign}(V_{IS} - V_{IS}^*) x_2 - \hat{F}_2(0, x_2) \quad (6.35)$$

where $u_T \in \mathbb{R}^{n-2 \times 1}$, $u_{IS} \in \mathbb{R}^{2 \times 1}$, $\hat{F}_1(x_1, x_2)$, and $\hat{F}_2(0, x_2)$ are vectors of continuous functions which represent (local) equivalent system dynamics,²⁶ $\Gamma_1 \in \mathbb{R}^{n-2 \times n-2}$ and $\Gamma_2 \in \mathbb{R}^{2 \times 2}$ are diagonal matrices with strictly positive real values, and V_{IS}^* is a positive constant which characterizes a desired magnitude of energy. Then:

1. All solutions of the controlled system (6.30) asymptotically reach a global invariant set Ω_{IS} included in the same subspace as x_2 and defined by²⁷ $\langle \overline{V_{IS} - V_{IS}^*} \rangle = 0$.
2. The energy of the controlled system converges to a neighborhood ε_{IS} of V_{IS}^* .

This result simply highlights that a representation of the system energy can be considered to define a sliding surface and, therefore that this energy can serve directly for stabilization purposes. Moreover, some technical remarks that need to be made are as follows:

- Control law (6.34) aims at driving the states of subvector x_1 to converge toward 0. In the case of a positive definite function²⁸ V_T , (6.34) simply corresponds to a high-gain control law which guarantees the asymptotic convergence of x_1 to zero, provided $\hat{F}_1(x_1, x_2)$ is upper bounded.
- As the role of (6.35) is of major importance for the control, the closed-loop system behavior greatly depends on $\hat{F}_2(0, x_2)$. In particular, a suitable design of $\hat{F}_2(0, x_2)$ may afford the system trajectory to converge towards either a fixed point or a (quasi-)periodic orbit²⁹ (see [38]).

²⁶Deduced from analysis of the actual system.

²⁷ $\langle \bullet \rangle$: averaged value of $\bullet$.

²⁸Such a condition can nevertheless be relaxed.

²⁹While keeping the system energy (quasi-)constant.

6.3 Hyperchaos Control

According to the Rössler definition [50], hyperchaotic systems exhibit more than one positive Lyapunov exponent (i.e., their dynamics can expand in more than one direction). Therefore, such systems are more sensitive to perturbations, external disturbances, and parameter variations than basic chaotic ones, leading to some particular difficulties for the control. In contrast, hyperchaotic systems provide a great “richness” in UPOs. Thus, such systems are highly attractive for some application fields such as chaos-based encryption, secured communication, etc. provided an efficient and suitable control of their potential behaviors can be achieved. Obviously, many methods coming from the control theory framework are able to deal with hyperchaotic systems. However, such methods often yield to closed-loop dynamics, far from the intrinsic characteristics of the original systems. In other words, most of these methods deal with hyperchaotic systems as with any nonlinear systems, without directly exploiting both their richness and properties. In contrast, nonconventional control techniques derived from the OGY method have been proposed within the literature of the field. As a result, these methods aim at directly exploiting the chaotic properties of the systems, for stabilization purposes, while preserving, as much as possible, their original characteristics. This section aims at presenting some of these methods, to bring into the focus some particularities and fundamental problems occurring when dealing with the control of hyperchaotic systems. In particular, this section highlights a modified OGY method sometimes referred to as the YLM method, and an enhanced version of this latest one, based on the use of the so-called Adaptive Adjustment Mechanism (AAM) [29].

6.3.1 The YLM Method for Hyperchaos Control

First, recall that OGY method aims at stabilizing an unstable orbit in the neighborhood of a hyperbolic fixed point by constraining the orbit to fall on the stable manifold. As the dynamics of hyperchaotic systems can expand in more than one direction, the basic OGY method then appears to be unsuitable to directly control such systems. The extended method proposed by Romeiras et al. [49] overcomes this problem, in part, with the use of a feedback matrix. However, this feedback control modifies the stability property of the fixed point by making this point to be fully stable after the parameter adjustment. To deal with a “highly” unstable fixed point,³⁰ while keeping some features of the hyperchaotic system, Yang et al. [67]

³⁰Fixed point at which at least two eigenvalues of the Jacobian matrix have modulus greater than unity.

proposed a control strategy which mainly consists in stabilizing only one of the unstable directions by means of time-dependent adjustment of control parameters. From a technical viewpoint, this method can be presented as follows.

First, note that the embedding theorem of Takens [59] asserts that if an orbit is in an attractor in the phase space, then the corresponding orbit in the embedding space is also in an attractor. Moreover, this theorem asserts that the two attractors have the same dimension. Then, with such assertions in mind, let us consider that the unstable orbit to be controlled is in an embedding space of dimension N , where N is a finite integer, and near a fixed point at which the dimension of the unstable manifold is N_u , with $N_u \leq N$.

Now, let us define a map $T : \zeta_n \rightarrow \zeta_{n+1}$ such that,

$$\zeta_{n+1} = F(\zeta_n, \varepsilon) \quad (6.36)$$

where $\zeta \in \mathbb{R}^N$ is the dynamical variable, $\varepsilon \in \mathbb{R}^{N_u}$ is a parameter vector assumed to be accessible for external adjustment, and $F(\zeta_n, \varepsilon)$ is a continuously differentiable, vector-valued function of ζ_n with ε as parameter.

In addition, let ζ_*^0 be the fixed point of the map (6.36) with $\varepsilon = 0$. Then, the purpose is to slightly adjust the parameter ε to control an orbit of the map that runs away from the (unstable) fixed point if $\varepsilon = 0$.

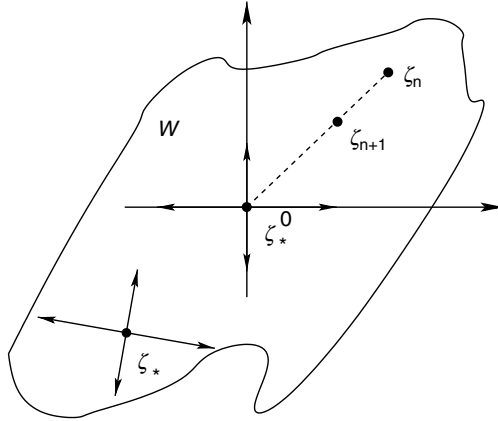
Moreover, following the same approach as for the OGY method, define the Jacobian matrix J of the map (with $\varepsilon = 0$) evaluated at the fixed point ζ_*^0 :

$$J = \left. \frac{\partial F_0}{\partial \zeta_n} \right|_{\zeta_n = \zeta_*^0} \quad (6.37)$$

where, for convenience, ζ_*^0 is assumed to be the origin of the N -dimensional space.

Then, the key point is that, according to the implicit function theorem, it can be asserted [67] that the map (6.36), with small parameters ε , has a fixed point, ζ_* , in a neighborhood of ζ_*^0 , provided the determinant of J is not equal to zero. In this case, for small ε values, one can consider that there exists a neighborhood W of ζ_*^0 that is large enough to also include a neighborhood of ζ_* , so that if ζ_n is in W then its image under the mapping (6.36) is also in W (see Figure 6.6). In this context, the control problem can be expressed as adjusting the parameter ε at each iteration so that ζ_{n+1} becomes closer to ζ_*^0 (located at the origin) than ζ_n (i.e., that $\zeta_{n+1} = k \zeta_n$, where k is a constant and $-1 < k < 1$).

In contrast to the OGY method which makes use of the stable submanifold to perform such convergence toward the fixed point, Yang et al. [67] proposed to drive the displacement of ζ_n on the Poincaré map surface by

**FIGURE 6.6**

Schematic representation of the YLM method.

means of a parameter adjustment of the form,

$$\varepsilon_n = M^{-1} (J - I)^{-1} (J - kI) \zeta_n \quad (6.38)$$

where I is the $N \times N$ identity matrix and M is the following matrix, assumed to be nonsingular.

$$M = \left. \frac{\partial \zeta_*}{\partial \varepsilon} \right|_{\varepsilon=0} \quad (6.39)$$

Such an adjustment of parameter ε , at each iteration, then leads the series $\{\zeta_1, \zeta_2, \dots\}$ to converge monotonically toward the fixed point ζ_*^0 , leading, finally, to the stabilization of the system.

REMARKS

This control strategy can deal with:

- Experimental applications for which a model of the system dynamics is not available, as matrices J and M can be defined experimentally [54]
- High-dimensional chaotic systems, as the method has been expressed for N -dimensional systems (N being a finite integer)
- Hyperchaotic systems with no pre-existing stable manifold

However, the YLM method is based on a parameter perturbations mechanism which requires at least one adjustable control parameter of the system to be found (and thus selected). Then, by adjusting this parameter, one of the unstable directions becomes stable so as to stabilize the whole unstable

orbit. However, in many experimental systems such as biological, chemical, or economical systems, the selection of such a parameter (*a priori*) is not trivial. To overcome this problem, Bu et al. [12] proposed a strategy based on an AAM [29], whose essence is discussed subsequently.

6.3.2 Enhanced YLM-Method with AAM

First, let us begin by presenting the Adaptive Adjustment Mechanism (AAM), as established by Huang [29].

6.3.2.1 The AAM Mechanism

Consider an n -dimensional nonlinear discrete system defined by

$$x(k+1) = F(x(k)) \quad (6.40)$$

where $x \in \mathbb{R}^n$ is the state vector and F is a vector-valued function.

Then, as proposed by Huang [29], the AAM considers a modified system of the form,

$$x(k+1) = \bar{F}(x(k)) = (1 - \gamma)F(x(k)) + \gamma x(k) \quad (6.41)$$

where γ is a positive control parameter referred to as an adaptive parameter.

REMARK 5

According to (6.41), the AAM forces a feedback adjustment whenever any variable strays away from its previous state.

The key point is that F and $\bar{F}$ share some properties which can be stated as follows [29].

COROLLARY 1

Systems F and $\bar{F}$ share exactly the same set of fixed points.

Moreover, by considering the eigenvalues of the respective Jacobian matrices $\bar{J}$ and J of both the original and modified systems evaluated at the same point, it can be seen that:

COROLLARY 2

For each and every fixed point of F and $\bar{F}$, there exists the following one-to-one correspondence between their eigenvalues:

$$\bar{\lambda}_j = (1 - \gamma)\lambda_j + \gamma \quad j = 1, 2, \dots, n$$

where $\bar{\lambda}_j$ and λ_j are the eigenvalues of $\bar{J}$ and J , respectively.

With such properties in mind, let us now consider the following result [29].

THEOREM 1

For an n -dimensional dynamical system defined by (6.40), supposing that x_ is a fixed point of F and a_j is the real part of the eigenvalues λ_j , for $j = 1, 2, \dots, n$ then:*

Case 1: if $a_j < 1$ for all $j = 1, 2, \dots, n$ there exists a $\bar{\gamma}$ such that, for all $\gamma \in (\bar{\gamma}, 1)$, all respective modulus under the AAM can be reduced to a magnitude that is less than unity, and hence, the original unstable fixed point will be stabilized.

Case 2: if $a_j > 1$ for all $j = 1, 2, \dots, n$ there exists a $\bar{\gamma}$ such that, for all $\gamma \in (1, \bar{\gamma})$, all respective modulus under the AAM can be reduced to a magnitude that is less than unity, and hence, the original unstable fixed point will be stabilized.

Case 3: if some a_j values are greater than unity, but others are less than unity, then the unstable fixed point cannot be stabilized by the simple AAM defined by (6.41).

The AAM provides a method to stabilize a broad class of multidimensional dynamical systems by considering one control parameter γ which depends on neither the structure of the original system nor any systems parameters. Then, with respect to these results, Bu et al. [12] investigated the problem of combining both the YLM method and the AAM, while retaining the interesting properties of each. From a technical viewpoint, in order to stabilize a fixed point x_f of (6.40), their result considers, with analogy to Equation (6.41), the following control strategy,

$$x_{k+1} = F(x_k) + M(F(x_f) - x_k) \quad (6.42)$$

where M is an $n \times n$ matrix to be defined.

For this purpose, consider a linear approximation of Equation (6.42) in a neighborhood W of the fixed point x_f ,

$$\delta x_{k+1} \approx Jx_k + (M - I) \delta x_k \quad (6.43)$$

where $J = (\partial F / \partial x_k)|_{x_k=x_f}$ is the Jacobian matrix of the original system evaluated at the fixed point x_f , I is the $n \times n$ identity matrix, and $\delta x_k = x_k - x_f$ is a small deviation of x_k from x_f .

Now, to solve the control problem which consists in finding a control law such that $\lim_{k \rightarrow \infty} \|\delta x_k\| \rightarrow 0$, Bu et al. [12] suggested to consider a relation of the form,

$$\delta x_k = \sigma(k - k_0) \delta x_{k_0} \quad (6.44)$$

where $\delta x_{k_0} = x_{k_0} - x_t$, $x_{k_0} \in W$ is the point from which the control is imposed on the original freely evolving system and $\sigma(k)$ is a scalar function which satisfies $\sigma(k) \rightarrow 0$ as $k \rightarrow \infty$.

Then, by assuming that $(J - I)$ is invertible, gathering (6.43) and (6.44) leads to define the matrix M as,

$$M = \left(\frac{\sigma(k+1)}{\sigma(k)} I - J \right) (J - I)^{-1} \quad (6.45)$$

Finally, let us express some technical and concluding remarks:

- With respect to Equation (6.45), there exists multiple possibilities to define $\sigma(k)$. For instance, as suggested earlier [12], $\sigma(k)$ can be of the form:

$$\sigma(k) = \gamma^k$$

where γ is a constant such that $\gamma \in [-1, 1]$.

- When compared with the YLM method, the AAM does not require *a priori* selection of a control parameter. Moreover, selecting of γ value (within the mentioned range) definitely determines the matrix M , which need not be changed afterwards with discrete time. Thus, in some senses, implementation of the control scheme can be quite straightforward.
- Finally, when compared with the original AAM, this method can be applied to a larger class of fixed points than the one restricted to hyperbolic points (see Theorem 1, Case 3).

6.4 Conclusions

This chapter is aimed at introducing the interest in dealing with the control of chaotic and hyperchaotic systems. Then, several control methodologies (which are by no means the only existing ones), coming from either the control theory framework or other seminal approaches, were presented. We hope that this presentation will help the readers in understanding the basics of the field of chaos control, and in appreciating some of the existing methods.

Acknowledgments

The author wishes to thank Professor J.P. Barbot for his helpful remarks and the great attention he paid on this chapter. The author is also grateful to Dr. M. Djemai for helpful discussions. Finally, the author wishes to thank Professor N.K. M'Sirdi for having given his support to this work.

References

1. E.H. Abed, H.O. Wang, and R.C. Chen, Stabilization of period doubling bifurcations and implications for control of chaos, *Physica D*, 70, 154–164, 1994.
2. F.T. Arecchi, G. Basti, S. Boccaletti, and A.L. Perrone, Adaptive recognition of a chaotic dynamics, *Europhys. Lett.*, 26, 327–332, 1994.
3. F.T. Arecchi, S. Boccaletti, M. Ciofini, R. Meucci, and C. Grebogy, The control of chaos: theoretical schemes and experimental realizations, *Int. J. Bifurcat. Chaos*, 8 (8), 1643–1655, 1998.
4. P.J. Aston and C.M. Bird, Analysis of the control of chaos — extending the basin of attraction, *Chaos Solitons Fract.*, 8, 1413–1429, 1997.
5. K. Aström and B. Wittenmark, *Adaptive Control*, 2nd ed., Addison-Wesley, 1995.
6. T. Basar and P. Bernhard, *H_∞-Optimal Control and Related Minimax Design Problems — A Dynamic Game Approach*, Systems and Control: Foundations and Applications, Birkhauser, 1991.
7. M. Basso, R. Genesio, and A. Tesi, Stabilizing periodic orbits of forced systems via generalized Pyragas controllers, *IEEE Trans. Circ. Syst. I*, 44, 1023–1027, 1997.
8. M. Basso, R. Genesio, and A. Tesi, Frequency domain methods and control of complex dynamics, in *Proceedings of the IEEE Conference on Decision and Control*, Florida, 1998, pp. 1936–1941.
9. S. Bhajekar, E.A., Jonckheere, and A. Hammad, H_{∞} control of chaos, in *Proceedings of the 33rd Conference Decision Control*, Lake Buena Vista, FL, 1994, pp. 3285–3286.
10. M.E. Bleich and J.E.S. Socolar, Stability of periodic orbits controlled by time-delay feedback, *Phys. Lett. A*, 210, 87–94, 1996.
11. S. Boccaletti and F.T. Arecchi, Adaptive control of chaos, *Europhys. Lett.*, 31, 127–132, 1995.
12. S. Bu, S. Wang, and H. Ye, Control of hyperchaos, *Phys. Rev. E*, 64 (5), p. 056212, 2001.
13. G. Chen and X. Dong, *From Order to Chaos: Methodologies, Perspectives and Applications*, World Scientific, Singapore, 1998.
14. C.L. Chen and W.Y. Lin, Sliding mode control for nonlinear systems with global invariance, *Proc. Inst. Mech. Eng.*, 211, 75–82, 1997.
15. G. Chin, L.R. Senesac, W.E. Blass, and J.J. Hillman, Stabilizing lead-salt diode lasers: understanding and controlling chaotic frequency emission, *Science*, 274, 1498, 1996.

16. M. Di Bernardo, An adaptive approach to the control and synchronization of continuous-time chaotic systems, *Int. J. Bifurcat. Chaos*, 6 (3), 557–568, 1996.
17. W.L. Ditto, S.N. Rauseo, and M.L. Spano, Experimental control of chaos, *Phys. Rev. Lett.*, 65 (26), 3211–3214, 1990.
18. W.L. Ditto, M.L. Spano, and J.F. Lindner, Techniques for the control of chaos, *Physica D*, 86 (1–2), 198–211, 1995.
19. W.L. Ditto and T. Munakata, Principles and applications of chaotic systems, *Commun. ACM*, 38 (11), 96–102, 1995.
20. X. Dong, G. Chen, and L. Chen, Adaptive control of the uncertain Duffing oscillator, *Int. J. Bifurcat. Chaos*, 7 (7), 1651–1658, 1997.
21. M. Fliess, Generalized controller canonical forms for linear and nonlinear dynamics, *IEEE Trans. Automat. Contr.*, AC-35, 994–1001, 1990.
22. A.L. Fradkov and J. Evans, Control of chaos: survey 1997–2000, in *Proceedings of the 15th IFAC World Congress on Automatic Control*, Barcelona, 2002, pp. 143–154.
23. A.L. Fradkov and A.Y. Pogromsky, Introduction to control of oscillations and chaos, *World Sci. Ser. Nonlinear Sci., Ser. A*, 35, 1998.
24. Z. Galias and M.J. Ogorzalek, Bang–bang control of chaotic systems, in *Proceedings of the International Workshop on Nonlinear Dynamics of Electronics Systems*, NDES'95, Dublin, 1995, pp. 229–232.
25. Z. Gills, C. Iwata, R. Roy, I.B. Swartz, and I. Triandaf, Tracking unstable steady states: extending the stability regime of a multimode laser system, *Phys. Rev. Lett.*, 79, 3169–3172, 1992.
26. G.A. González, Controlling chaos of an uncertain Lozi system via adaptive techniques, *Int. J. Bifurcat. Chaos*, 5, 559–562, 1994.
27. G.A. González, M. Tropicarsky, and C.E. D'Attellis, A remark on chaotic behavior in adaptive control systems, *IEEE Trans. Automat. Contr.*, 39, 2145–2148, 1994.
28. C. Grebogy and Y.C. Lai, Controlling chaos in high dimensions, *IEEE Trans. Circ. Syst. I*, 44, 971–975, 1997.
29. W. Huang, Stabilizing nonlinear dynamical systems by an adaptive adjustment mechanism, *Phys. Rev. E*, 61 (2), R1012–R1015, 2000.
30. A.W. Hübler, Adaptive control of chaotic systems, *Helvetica Phys. Acta*, 62, 343–346, 1989.
31. E.R. Hunt, Stabilizing high-periodic orbits in a chaotic system: the diode resonator, *Phys. Rev. Lett.*, 67 (13), 1953–1955, 1991.
32. S. Iplikci and Y. Denizhan, Targeting in dissipative chaotic systems: a survey, *Chaos*, 12 (4), 995–1005, 2002.
33. M.J. Jang, C.L. Chen, and C.K. Chen, Sliding mode control of chaos in the cubic Chua's circuit system, *Int. J. of Bif. and Chaos*, 12 (6), 1437–1449, 2002.
34. E.A. Jonckheere, A. Hammad, C.-Y. Cheng, and C.-C. Chien, LQ control of chaos, in *Proceedings of the IFAC Symposium on Robust Control Design*, 1994.
35. W. Just, T. Bernard, M. Ostheimer, E. Reibold, and H. Benner, Mechanism of time-delayed feedback control, *Phys. Rev. Lett. A*, 78, 203–206, 1997.
36. T. Kapitaniak, *Controlling Chaos*, Academic Press, New York, 1996.
37. L. Laval and N.K. M'Sirdi, Controlling chaotic and hyperchaotic systems via energy regulation, *Chaos Solitons Fract.*, 15, 285–295, 2003.
38. L. Laval and N.K. M'Sirdi, On stabilizing n-dimensional chaotic systems, *Int. J. Bifurcat. Chaos*, 13 (2), 473–481, 2003.

39. T.L. Liao and S.H. Lin, Adaptive control and synchronization of Chua's circuit, *Asian J. Contr.*, 1 (2), 75–87, 1999.
40. J.F. Lindner and W.L. Ditto, Removal suppression and control of chaos by nonlinear design, *Appl. Mech. Rev.*, 48 (12), 795–807, 1995.
41. H. Nakajima and Y. Ueda, Limitation of generalized delayed feedback control, *Physica D*, 111, 143–150, 1998.
42. T.C. Newell, P.M. Alsing, A. Gavriedlides, and V. Kovanis, Synchronization of chaotic diode resonators by occasional proportional feedback, *Phys. Rev. Lett.*, 72, 1647, 1994.
43. E. Ott, C. Grebogy, and J.A. Yorke, Controlling chaos, *Phys. Rev. Lett.*, 64, 1196–1199, 1990.
44. N.H. Packard, J.P. Crutchfield, J.D. Farmer, and R.S. Shaw, Geometry from time series, *Phys. Rev. Lett.*, 45, 712, 1980.
45. B. Peng, V. Petrov, and K. Showalter, Controlling chemical chaos, *J. Phys. Chem.*, 95, 4957, 1991.
46. I. Prigogine and I. Stengers, *Order Out of Chaos: Man's New Dialogue with Nature*, Bantam Books, New York, 1988.
47. K. Pyragas, Continuous control of chaos by self-controlling feedback, *Phys. Lett. A*, 170, 421–428, 1992.
48. K. Pyragas, Control of chaos via an unstable delayed feedback controller, *Phys. Rev. Lett.*, 86 (11), 2265–2268, 2001.
49. F.J. Romeiras, C. Grebogy, E. Ott, and W.P. Dayawansa, Controlling chaotic dynamical systems, *Physica D*, 58, 165–192, 1992.
50. O.E. Rössler, An equation for hyperchaos, *Phys. Lett.*, A71 (2–3), 155–157, 1979.
51. R. Roy, T.W. Murphy Jr., T.D. Mayer, Z. Gillis, and E.R. Hunt, Dynamical control of a chaotic laser: experimental stabilization of a globally coupled system, *Phys. Rev. Lett.*, 68, 1259–1262, 1992.
52. R.M. Sanner and J.J. Slotine, Gaussian networks for direct adaptive control, *IEEE Trans. Neural Networks*, 3, 837–863, 1992.
53. T. Sauer and J.A. Yorke, How many delay coordinates do you need, *Int. J. Bifurcat. Chaos*, 3, 737, 1993.
54. T. Sauer and J.A. Yorke, Reconstructing the Jacobian from data with observational noise, *Phys. Rev. Lett.*, 83, 1331–1334, 1999.
55. J.J. Slotine and W. Li, *Applied Nonlinear Control*, Prentice-Hall International Limited, London, 1991.
56. J.E.S. Socolar, D.W. Sukow, and D.J. Gauthier, Stabilizing unstable periodic orbits in fast dynamical systems, *Phys. Rev. E*, 50, 3245–3248, 1994.
57. T. Shinbrot, C. Grebogi, E. Ott, and J.A. Yorke, Using small perturbations to control chaos, *Nature*, Vol. 363, pp. 411–417, 1993.
58. S.H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*, Addison-Wesley, 1994.
59. F. Takens, Detecting strange attractors in turbulence, in *Lecture Notes in Mathematics, Dynamical Systems and Turbulence*, Vol. 898, Springer, Berlin, 1981, pp. 366.
60. A. Tsonis, *Chaos: From Theory to Applications*, Plenum Press, New York, 1992.
61. T. Ushio, Limitation of delayed feedback control in nonlinear discrete-time systems, *IEEE Trans. Circ. Syst. I*, 43, 815–816, 1996.

62. V.I. Utkin, Variable structure systems with sliding mode: a survey, *IEEE Trans. Automat. Contr.*, ACC-22 (2), 212–222, 1977.
63. V.I. Utkin, *Sliding Modes in Control and Optimization*, Springer-Verlag, 1992.
64. H.T. Yau, , C.K. Chen, and C.L. Chen, Sliding mode control of chaotic systems with uncertainties, *Int. J. Bifurcat. Chaos*, 10 (5), 1139–1147, 2000.
65. T.L. Vincent and J. Yu, Control of a chaotic system, *Dyn. Contr.*, 1, 35–52, 1991.
66. T.H. Yang, S.F. Chen, and Y.S. Gou, Efficient strategy for the occasionally proportional feedback method in controlling chaos, *Phys. Rev. E*, 59 (5), 5393–5399, 1999.
67. L. Yang, Z. Liu, and J-M Mao, Controlling hyperchaos, *Phys. Rev. Lett.*, 84 (1), 60–70, 2000.
68. X. Yu, Controlling Lorenz chaos, *Int. J. Syst. Sci.*, 27 (4), 355–359, 1996.
69. X. Yu, G. Chen, Y. Song, Z. Cao, and Y. Xia, A generalized method for controlling higher order chaotic systems, in *Proceedings of the IEEE Conference on Decision Control*, Sydney, Australia, 2000, pp. 2054–2059.
70. X. Yu, G. Chen, Y. Xia, Y. Song, and Z. Cao, An invariant-manifold-based method for chaos control, *IEEE Trans. Circ. Syst. I: Fundam. Theory Appl.*, 48 (8), 930–937, 2001.
71. H. Zhang, H. Qin, and G. Chen, Adaptive control of chaotic systems with uncertainties, *Int. J. Bifurcat. Chaos*, 8 (10), 2041–2046, 1998.
72. K. Zhou, J.C. Doyle, and K. Glover, “Robust and Optimal Control”, Prentice-Hall, 1995.

Polytopic Observers for Synchronization of Chaotic Maps

G. Millérioux and J. Daafouz

CONTENTS

7.1	Introduction	324
7.2	Preliminaries	325
7.2.1	Chaotic Systems with Polytopic Description	325
7.2.1.1	Lur'e Systems	326
7.2.1.2	Output Injection with Time-Varying Dynamical Matrix	326
7.2.1.3	Piecewise Linear Systems	327
7.2.2	Poly-Quadratic Stability	327
7.3	Message-Free Chaos Synchronization	331
7.3.1	Synchronization and State Reconstruction	331
7.3.2	Polytopic Observers	332
7.3.3	Conditions of Global Synchronization	333
7.4	Message-Embedded Chaos Synchronization	334
7.4.1	Input Independent Global Synchronization	334
7.4.2	Polytopic Unknown Input Observers	335
7.4.3	Conditions of IIGS	336
7.4.4	Plaintext Recovering	337
7.5	Illustrative Examples	338
7.5.1	Unknown Input Observer Design	338
7.5.2	Real-Time Private Communication Experiment	338
7.6	Conclusion	341
	References	341

7.1 Introduction

The interest in chaos synchronization has increased in the last decade since the pioneering works by Pecora and Carroll [30, 31]. Message-free chaos synchronization has entered the control scene and has become a popular open problem from the control theory point of view [2]. Attempts [3] have been made to give a general formalism for synchronization in dynamical systems and many special issues devoted to the subject are of particular interest [38–41]. A wide variety of methods have been investigated to achieve synchronization of two dynamical systems coupled in a unidirectional way. A unidirectional coupling involves a system called *response* forced by an external signal emanating from a system called *drive* which exhibits a chaotic behavior. From the control theory point of view, three main suitable approaches are of interest. The first one consists of the reconstruction of the attractors from a sliding window of a finite amount of output measurements from the chaotic system [17, 36]. Such a method is motivated by the Takens' theorem [42]. The second approach is referred to as controlled synchronization and consists of finding closed loop feedback control to ensure synchronization. This requires the measure of all the state variables of the system. Finally, when only partial information of these variables is available, meaning that only output variables of the drive are transmitted to the response, observer-based methods can be considered. One of the important surveys on chaos synchronization dealing with the observer approach is presented in Nijmeijer and Mareels [28]. For relatively recent results, the reader can refer to Huijberts et al. [16] for observers with linearizable dynamics, Pogromsky and Nijmeijer [32] for observers derived from the concept of absolute stability, Ramirez and Hernandez [37] for observers dedicated to systems having a generalized Hamiltonian forms, and Millerioux and Daafouz [23, 24] for observers of systems having polytopic description and whose design is based upon linear matrix inequalities (LMI).

On the other hand, one of the well-known practical interests of chaos synchronization lies in the potential applications in communications, and more specifically in the possibilities of encoding or masking messages by suitable embedding. Indeed, it is reasonable to think that there is likely a connection between the random-like behaviors exhibited by chaotic systems and the required properties like confusion and diffusion of cryptosystems. Chaos-based encryption is currently an active field of research. A survey of chaos-based encryption schemes with an adequate bibliography incorporating some cryptographic skills has been discussed earlier [18, 19]. In Ref. [5, 29, 44], an overview of the techniques currently relevant for transmitting information via a chaotic signal is given. As far as terminology is concerned, in a message-embedded context, the system exhibiting

chaos is commonly referred to as the *transmitter*, whereas the system which must extract the information is named the *receiver*.

In this chapter, both message-free and message-embedded chaos synchronization problems are investigated over a unified framework which involves polytopic observers. Indeed, based upon the fact that the chaos generator named drive in the message-free context or transmitter in the message-embedded context exhibits a chaotic motion, its underlying state variables are constrained to a compact domain. As a consequence, the range and the bounds of those variables can be known such that the dynamical matrices can be expressed in a polytopic way. Then, the error of synchronization can also be written in a polytopic form. By using the notion of poly-quadratic stability [6, 7] and a parameter dependent Liapunov function (PDLF), the problem of the observer synthesis can be turned into the resolution of a LMI set. The benefits of such an approach rely on the fact that the resulting synchronization is global, and the computation of the gains of the observer is carried out in a systematic and tractable way.

The layout of the chapter is the following. In Section 7.2, chaotic systems admitting a polytopic description are presented. Then, some background concerning the notion of poly-quadratic stability are presented. Section 7.3 and Section 7.4, respectively, deal with the message-free and the message-embedded chaos synchronization problem. A systematic procedure is stated for the design of the polytopic observers which must achieve global synchronization in the message-free context and additionally the recovering of the masked information in the message-embedded context.

Notation Throughout this chapter, $\mathbf{1}_n$ is the n -dimensional identity matrix and $\mathbf{0}_{n \times m}$ the $n \times m$ null matrix. For a matrix X , X^T stands for its transpose. When symmetric, $X > 0$ indicates that X is positive definite. $X^\dagger$ corresponds to the Moore–Penrose generalized inverse of X given by $X^\dagger = (X^T X)^{-1} X^T$.

7.2 Preliminaries

7.2.1 Chaotic Systems with Polytopic Description

In this section, the class of chaotic systems under consideration is detailed. We concentrate on dynamical systems of which general description is as follows:

$$\begin{cases} x_{k+1} = A(\rho_k)x_k + E(\rho'_k) \\ y_k = Cx_k \end{cases} \quad (7.1)$$

where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$, $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{m \times n}$. Let ρ and ρ' be, respectively, an L -dimensional and M -dimensional function. The images under ρ and ρ' are, respectively, denoted $\rho_k = (\rho_k^1, \dots, \rho_k^L)^T$ and $\rho'_k = (\rho_k'^1, \dots, \rho_k'^M)^T$. They are both assumed to be available. The term $A(\rho_k)x_k$ includes all the affine dependence in the dynamics with respect to x_k . Besides, A is of class C^1 with respect to the entries of ρ_k such that A can be rewritten in the form $A(\rho_k) = A_0 + \sum_{i=1}^L \rho_k^i A_i^{lc}$. A_0 is the matrix derived from $A(\rho_k)$ by keeping its constant entries while setting to zero its time-varying entries. A_i^{lc} is a matrix whose entries are all zero except the one, located at line l and column c , which equals unity. The superscripts l and c depend on i and correspond, respectively, to the position of ρ_k^i in $A(\rho_k)$. E is a (possible) nonlinear n -dimensional function depending on ρ'_k (not necessarily in an affine way).

The quantity ρ_k acts as a time-varying parameter for A and is assumed to be bounded such that A lies in a compact set which may be embedded in a polytope, that is:

$$A(\rho_k) = \sum_{i=1}^N \xi_k^i(\rho_k) A_i \quad (7.2)$$

The A_i 's correspond to the vertices of the convex hull $\text{Co}\{A_1, \dots, A_N\}$ and the ξ_k^i 's belong to the compact set $\mathcal{S} = \{\mu_k \in \mathbb{R}^N, \mu_k = (\mu_k^1, \dots, \mu_k^N)^T, \mu_k^i \geq 0 \forall i, \text{ and } \sum_{i=1}^N \mu_k^i = 1\}$. The ξ_k^i 's can always be expressed as a linear function of the ρ_k^i 's. The class of systems described by (7.1) includes some usual chaotic systems.

7.2.1.1 Lur'e Systems

These systems are described by the discrete model:

$$x_{k+1} = A_1 x_k + E(y_k)$$

A_1 is a constant dynamical matrix and E is an n -dimensional function of the output y_k . Such systems are derived from (7.1) by letting $A(\rho_k) = A_1$ and $\rho'_k = y_k$. Besides, note that a constant matrix A_1 is a special case of (7.2) with $N = 1$.

7.2.1.2 Output Injection with Time-Varying Dynamical Matrix

These systems are described by the recursion:

$$x_{k+1} = A(y_k)x_k + E(y_k)$$

Whenever such systems exhibit a chaotic behavior, x_k and, therefore, y_k are constrained to a compact domain. As a consequence, y_k is bounded

at least in a hypercube because each component y_k^i ranges between two extremal values $\underline{y}_k^i$ and $\bar{y}_k^i$. Hence, $A(y_k)$ can be expressed in a polytopic way and verifies (7.2) by letting $\rho_k = \rho'_k = g(y_k)$, g being a function of y_k , such that A is of class C^1 with respect to ρ_k . A polytopic decomposition for such systems is detailed in Ref. [24, 25].

7.2.1.3 Piecewise Linear Systems

These systems are described by the recursion:

$$x_{k+1} = A_i x_k + E_i$$

Let the state space $\mathbb{R}^n$ be partitioned into N distinct regions R_i , with $\bigcup_{i=1}^N R_i \subseteq \mathbb{R}^n$. A_i and E_i are constant matrices assigned, with a one-to-one correspondence, to the region R_i , visited by x_k at the discrete time k . Piecewise linear systems are derived from (7.1) as the following. Let ρ (resp. ρ') be a scalar function $\mathbb{R}^n \rightarrow I$, with $I = \{1, \dots, N\}$ an index set of N elements, defined by $\rho(x_k) = \rho_k = i$ (resp. $\rho'(x_k) = \rho'_k = i$), if x_k visits the region R_i at the discrete time k . Let $\rho'_k = \rho_k$ and parameterizing $A(\rho_k)$ (resp. $E(\rho_k)$) such that $A(\rho_k) = A_i$ (resp. $E(\rho_k) = E_i$) when $\rho_k = i$, the usual piecewise linear description mentioned previously is thus obtained. Now, defining an indicator vector $\xi_k = (\xi_k^1, \dots, \xi_k^N)^T$ as follows:

$$\xi_k^i = \begin{cases} 1 & \text{if } \rho_k = i \\ 0 & \text{otherwise} \end{cases}$$

Thus, $A(\rho_k)$ can be expressed in the polytopic form (7.2).

7.2.2 Poly-Quadratic Stability

Poly-quadratic stability has been introduced earlier [6] to state necessary and sufficient conditions of existence of PDLFs in the context of linear parameter varying systems (LPV). The notion of LPV systems was first introduced by Shamma and Athans [34]. This class of systems is different from standard linear time-varying systems counterpart because of the dependence of the system matrices on the variations of the plant dynamics. The study of LPV systems was first motivated by the gain scheduling control design methodology, where the design of the controller involves the design of several linear time invariant controllers for a parameterized family of linearized models of a system and the interpolation of the controller gains [33, 34]. Although it seems to be working well in practice, this heuristic design procedure does not take the parameter variations into account

and cannot provide any stability or performance guarantees, except for slow varying parameters [35]. Also, the LPV theory has been found useful to simplify the interpolation and realization problems associated with conventional gain scheduling. Specially, it allows to treat gain scheduling controllers as a single entity, with the gain scheduling achieved entirely by the parameter-dependent controller [43]. An LPV system is given by:

$$v_{k+1} = \mathcal{A}(\xi_k)v_k \quad (7.3)$$

where v_k is an n -dimensional state vector and $\xi_k \in \Xi \subset \mathbb{R}^N$ is a bounded time-varying parameter. This system is referred as polytopic when the dependence of the dynamical matrix on ξ_k is given by:

$$\mathcal{A}(\xi_k) = \sum_{i=1}^N \xi_k^i A_i, \quad \xi_k = (\xi_k^1, \dots, \xi_k^N)^T, \quad \xi_k^i \geq 0, \quad \sum_{i=1}^N \xi_k^i = 1 \quad (7.4)$$

where A_i are constant matrices called vertices. What is known in the literature as quadratic stability refers to the checking stability of the previous system using classical quadratic Liapunov function $V(v_k) = v_k^T P v_k$, with P a positive definite matrix. Although any quadratic stability-based condition seems numerically useful because it generally leads to LMI feasibility problems [1], this kind of conditions is conservative. To reduce such a conservatism, PDLFs have been introduced [12]. This consists in letting the Liapunov matrix P depend on the parameter vector ξ_k . A general result is given in Theorem 1.

THEOREM 1

System (7.3) is asymptotically stable if there exists a Liapunov function

$$V(v_k, \xi_k) = v_k^T \mathcal{P}_k(\xi_k) v_k$$

such that

$$\alpha_1(\|v_k\|) \leq V(v_k, \xi_k) \leq \alpha_2(\|v_k\|) \quad (7.5)$$

and whose difference along the solution of (7.3) is negative definite descrescent, that is

$$\mathcal{L} = V(v_{k+1}, \xi_{k+1}) - V(v_k, \xi_k) \leq -\alpha_0(\|v_k\|) \quad (7.6)$$

for all $v_k \in \mathbb{R}^n$ and $\xi_k \in \Xi$ and where $\alpha_0(\cdot)$, $\alpha_1(\cdot)$, and $\alpha_2(\cdot)$ are κ_∞ functions.¹

¹A function $\alpha: [0, \infty) \rightarrow [0, \infty)$ is a κ_∞ function if it is continuous, strictly increasing, zero at zero, and unbounded ($\alpha(s) \rightarrow \infty$ as $s \rightarrow \infty$).

In practice, this result is very general and cannot be used in its present form because there is no systematic way to build the Liapunov function $V(v_k, \xi_k)$ as a function of the time-varying parameter ξ_k . On the basis of a similar description of the LPV system (7.3), poly-quadratic stability has been defined in order to check stability of the polytopic LPV system (7.3), with a PDLF of the form:

$$V(v_k, \xi_k) = v_k^T \mathcal{P}_k v_k \quad \text{with} \quad \mathcal{P}_k = \sum_{i=1}^N \xi_k^i P_i, \quad \xi_k^i \geq 0, \quad \sum_{i=1}^N \xi_k^i = 1 \quad (7.7)$$

where P_i are symmetric positive definite constant matrices of appropriate dimension. Such a PDLF satisfies Condition (7.5) with $\alpha_2(\|v_k\|) = \sum_{i=1}^N \lambda_{\max}(P_i) \|v_k\|^2$ and $\alpha_1(\|v_k\|) = \varepsilon \|v_k\|^2$, with ε a sufficiently small positive scalar. It's difference along the solution of (7.3) is given by

$$\mathcal{L} = V(v_{k+1}, \xi_{k+1}) - V(v_k, \xi_k) \quad (7.8)$$

with

$$V(v_{k+1}, \xi_{k+1}) = v_{k+1}^T \mathcal{P}_{k+1} v_{k+1}, \quad \mathcal{P}_{k+1} = \sum_{i=1}^N \xi_{k+1}^i P_i \quad (7.9)$$

A necessary and sufficient condition of existence of such a PDLF is proposed in [6]. It consists in checking the feasibility of a set of LMI in which the unknowns are directly related to the Liapunov matrices P_i . This condition allows to answer either “yes” or “no” to the following question: Is there a Liapunov function of the form (7.7) allowing to check that the LPV system (7.3) is globally asymptotically stable. Before stating this result, a definition of poly-quadratic stability is recalled from Ref. [6].

DEFINITION 1 System (7.3) is said to be poly-quadratically stable, if there exists a positive definite and quadratic PDLF V as defined in (7.7) whose difference along the solution of (7.3) satisfies

$$V(v_{k+1}, \xi_{k+1}) - V(v_k, \xi_k) = v_k^T (\mathcal{A}^T \mathcal{P}_{k+1} \mathcal{A} - \mathcal{P}_k) v_k < -\alpha_0(\|v_k\|) \quad (7.10)$$

with α_0 a κ_∞ function.

The following theorem gives a necessary and sufficient condition for the dynamics (7.3) to be poly-quadratically stable and so for v_k to converge globally toward zero.

THEOREM 2

The LPV system (7.3) is poly-quadratically stable if and only if there exist N positive symmetric matrices $P_1, \dots, P_N$ and N matrices $G_1, \dots, G_N$ satisfying the following set of LMI:

$$\begin{bmatrix} P_i & A_i^T G_i^T \\ G_i A_i & G_i + G_i^T - P_j \end{bmatrix} > 0 \quad (7.11)$$

for all $(i, j) \in \{1, \dots, N\} \times \{1, \dots, N\}$. The Liapunov function is then given by

$$V(v_k, \xi_k) = v_k^T \left(\sum_{i=1}^N \xi_k^i P_i \right) v_k$$

The main advantage in using poly-quadratic stability to answer the basic problem stated in the beginning of this section relies on the fact that it provides a sufficient condition of asymptotic stability with the following features:

- This condition is numerically well tractable. One has to check the feasibility of a set of LMI. This reduces to a convex optimization problem for which powerful numerical algorithms and routines are known to exist.
- It is obviously less conservative than the conditions based on checking the existence of a common Liapunov function.
- This condition can be used for switched linear systems as explained in Daafouz et al. [8]. When compared with some existing conditions for stability of switched linear systems which require that the matrices A_i commute for each i , the conservatism is reduced.
- Under the arbitrary switching rule, the proposed condition has to be satisfied for all the pairs (i, j) , that is, to take into account all possible switches from each subsystem to another. However, if all the transitions are not allowed and if one is able to determine a set of all ordered pairs (i, j) of indices denoting the possible switches from a subsystem A_i to another subsystem A_j , the proposed condition can be modified to take into account only these selected pairs of indices. Hence, knowledge of allowed transitions between subsystems is a way to reduce again the conservatism in the case of switched linear systems.
- The extra matrices G_i can be useful when design problems are formulated using this condition. The control or observer gains will depend on this matrices and not on the "Liapunov" matrices. Hence, if other constraints than stability are fixed, they will not affect the Liapunov

matrices but only these extra matrices [26]. This can be less conservative than the results where the control or the observer gains depend explicitly on the Liapunov matrices (see also Daafouz et al. [9] where the switched static output feedback problem is presented as an example of this situation).

In the following sections, the problems of message-free and message-embedded chaos synchronization are discussed. A motivation of the observer-based approach developed in both cases is previously given.

7.3 Message-Free Chaos Synchronization

7.3.1 Synchronization and State Reconstruction

Observer-based synchronization is one of the possible methods which enables the dynamical system state reconstruction. It is well known that for a single input–single output autonomous linear dynamical system of dimension n , a brute method, in view of reconstructing the state vector x_k at the discrete time k , consists in simply inverting a so-called observability matrix Q_0 . It stems from the fact that, for a realization (A, C) , x_k can be expressed as a linear combination of the n past outputs y_{k-i} , where $i = 0, \dots, n-1$ in the form:

$$x_k = A^{n-1}Q_0^{-1}Y_k \text{ with } Q_0 = \begin{bmatrix} C \\ \vdots \\ CA^{n-1} \end{bmatrix} \text{ and } Y_k = [y_{k-n+1}, \dots, y_k]^T$$

provided that Q_0 is invertible or equivalently that the pair (A, C) is observable. A state reconstruction based on past outputs has a counterpart for nonlinear systems given by the Taken's theorem [42].

THEOREM 3

Let Ω be a compact manifold of dimension n . Let φ be a smooth (at least of class C^2) diffeomorphism $\mathbb{R}^n \rightarrow \mathbb{R}^n$ and h a smooth function $\mathbb{R}^n \rightarrow \mathbb{R}$. Generically, there exists a map ϕ from $\mathbb{R}^n$ to $\mathbb{R}^{l+1}$ such that:

$$\phi(x_k) = (h(x_k), \dots, h(\varphi^l(x_k)))$$

for $l \geq 2n$.

Applying such a theorem would enable to reconstruct the state vector x_k of a chaotic map from at least the $2n + 1$ past output values. Unfortunately,

Taken's theorem does not provide the function ϕ . Besides, it is well known that a direct reconstruction, or equivalently a direct "inversion," suffers from measurement errors and is not viable in practice. It is better to use asymptotical reconstruction, an approach which motivates the use of observers. The issue of direct reconstruction has been discussed earlier [15, 17, 36].

Before dealing with polytopic observers, a definition of synchronization is presented.

DEFINITION 2 Global synchronization of a drive–response system can be expressed in one of the following formulations involving the respective state vectors x_k and $\hat{x}_k$:

$$\lim_{k \rightarrow \infty} \|x_k - \hat{x}_k\| = 0 \quad \forall \hat{x}_0 \quad (7.12)$$

$$\exists k_f, \quad \|x_k - \hat{x}_k\| = 0 \quad \forall \hat{x}_0 \text{ and } \forall k > k_f \quad (7.13)$$

Equation (7.12) corresponds to an asymptotical synchronization, whereas Equation (7.13) corresponds to a finite time synchronization.

7.3.2 Polytopic Observers

Consider a response system governed by the dynamics of the following observer named polytopic observer:

$$\begin{cases} \hat{x}_{k+1} = A(\rho_k)\hat{x}_k + E(\rho'_k) + L(\rho_k)(y_k - \hat{y}_k) \\ \hat{y}_k = C\hat{x}_k \end{cases} \quad (7.14)$$

The equation of the synchronization error between the drive and the response is obtained by subtracting (7.14) from (7.1).

$$\varepsilon_{k+1} = (A(\rho_k) - L(\rho_k)C)\varepsilon_k \quad (7.15)$$

It is recalled that $A(\rho_k)$ is written in a polytopic form (7.2) as described in Section 7.2.1. Let us define $L(\rho_k)$ as a time-varying gain matrix depending on ρ_k and satisfying the relation:

$$L(\rho_k) = \sum_{i=1}^N \xi_k^i(\rho_k) L_i \quad (7.16)$$

The terms ξ_k^i are the ones involved in the polytopic decomposition (7.2) of $A(\rho_k)$ and so depend on ρ_k . Equation (7.16) means that the gain matrix

$L(\rho_k)$ is interpolated from the vertices L_i . They have to be computed in order to ensure global synchronization. From this perspective, substituting (7.2) and (7.16) into (7.15) yields:

$$\varepsilon_{k+1} = \sum_{i=1}^N \xi_k^i (A_i - L_i C) \varepsilon_k \quad (7.17)$$

This is a polytopic formulation of the error of synchronization. Section 7.3.3 is devoted to conditions for global convergence conditions of (7.17) and so global synchronization of (7.1)–(7.14). They are based upon the notion of poly-quadratic stability presented in Section 7.2.2.

7.3.3 Conditions of Global Synchronization

PROPOSITION 1

Global synchronization of (7.1)–(7.14) is achieved whenever the feasibility condition of the following set of LMI is satisfied:

$$\begin{bmatrix} P_i & A_i^T G_i^T - C^T F_i^T \\ G_i A_i - F_i C & G_i + G_i^T - P_j \end{bmatrix} > 0 \quad (7.18)$$

for all $(i, j) \in \{1, \dots, N\} \times \{1, \dots, N\}$.

The G_i 's, P_i 's, and F_i 's are unknown matrices of appropriate dimensions. The resulting gains L_i are given by $L_i = G_i^{-1} F_i$.

REMARK 1

The formulation (7.19) differs from the one encountered in [7] but is strictly equivalent.

7.4 Message-Embedded Chaos Synchronization

7.4.1 Input Independent Global Synchronization

We concentrate on the message-embedded scheme depicted in Figure 7.1. The plaintext m_k is encrypted by an encryption function e which depends on a chaotic key stream x_k generated by the chaos generator named in this context the transmitter. The resulting quantity denoted u_k is embedded in the dynamics of the chaos generator. In the sequel, u_k will be abusively

called the information. It is important to note that u_k is not transmitted to the receiver, only an output signal y_k , whose dimension is less than the dimension of x_k , is transmitted through the channel to the receiver. The receiver system must be designed such that u_k can be recovered, given the only available data y_k . Once u_k is recovered, the plaintext m_k can be extracted by applying the decoding function e^{-1} , provided that $\hat{x}_k$ is exactly synchronized with x_k . This requirement is the main problem to be overcome.

Few methods involving observer-based approaches have been investigated. The design of the observer is based on the consideration of the state reconstruction error dynamics. On one hand, the design can be carried such that its convergence toward zero is guaranteed despite of the presence in the dynamics of a residual term involving u_k . A sufficient condition for the exponential convergence of the state-error dynamics in the case of Lipschitz nonlinearities is given in Ref. [21]. For discrete-time systems, dead-beat observers can be designed to make the matrix involved in the synchronization error equation to be nilpotent [11, 22]. In this case, despite the fact that the error depends on u_k , it can reach zero after a finite number of steps. This method works, for example, by performing a zero eigenvalues assignment with discrete-time Lur'e chaotic systems. On the other hand, the design can be performed in such a way that the residual term involving u_k no longer appears in the state reconstruction error dynamics. An example is presented in the context of a chaotic inverse system encryption approach with Lur'e systems [46]. The problem has been formally tackled by introducing the notion of input independent global synchronization (IIGS) [26, 27]. In Ref. [27], an unknown input observer approach dedicated to polytopic systems has been established. Such an approach is recalled here.

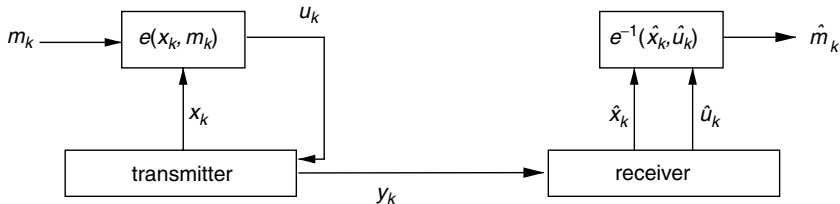


FIGURE 7.1

Message-embedded scheme.

DEFINITION 3 IIGS of a transmitter–receiver system can be expressed in one of the following formulations involving the respective state vectors x_k and $\hat{x}_k$:

$$\lim_{k \rightarrow \infty} \|x_k - \hat{x}_k\| = 0 \quad \forall \hat{x}_0, u_k \quad (7.19)$$

$$\exists k_f, \quad \|x_k - \hat{x}_k\| = 0 \quad \forall \hat{x}_0, \quad \forall k > k_f \text{ and } \forall u_k \quad (7.20)$$

Equation (7.20) and Equation (7.21) differ from Definition 2 by the fact that $\hat{x}_k$ must coincide with x_k given any u_k .

7.4.2 Polytopic Unknown Input Observers

To achieve an IIGS, an unknown input observer approach is presented. Unknown input observers have been largely investigated for linear systems [4, 10, 13, 45] or bilinear systems [20]. Extension of such a structure for systems having a polytopic description is presented [27].

Let us consider the transmitter described by:

$$\begin{cases} x_{k+1} = A(\rho_k)x_k + E(\rho'_k) + Bu_k \\ y_k = Cx_k \end{cases} \quad (7.21)$$

which differs from (7.1) by the introduction of an additional term Bu_k involving a so-called “mixing matrix” $B \in \mathbb{R}^{n \times r}$ ($r \geq m$) and $u_k \in \mathbb{R}^r$. As previously described in Section 7.4.1, the term u_k results from an encoding function e such that $u_k = e(x_k, m_k)$. It is assumed that e admits an inverse decoding function e^{-1} such that $m_k = e^{-1}(x_k, u_k)$. As far as the receiver part is concerned, a natural extension of the linear unknown input observer structure is proposed:

$$\hat{x}_{k+1} = (PA(\rho_k) - L(\rho_k)C)\hat{x}_k + PE(\rho'_k) + L(\rho_k)y_k + Qy_{k+1} \quad (7.22)$$

with $P = \mathbf{1}_n - QC$ and $L(\rho_k) = \sum_{i=1}^N \xi_k^i(\rho_k)L_i$.

The gains Q and L_i 's ($i = 1, \dots, N$) must be computed to achieve an IIGS of (7.22) and (7.23) and, consequently, the global convergence of the state reconstruction error ε_k . Its dynamics is obtained by subtracting (7.23) from (7.22) and taking into account the polytopic expression of $L(\rho_k)$ and $A(\rho_k)$:

$$\varepsilon_{k+1} = \sum_{i=1}^N \xi_k^i(PA_i - L_iC)\varepsilon_k + PBu_k \quad (7.23)$$

7.4.3 Conditions of IIGS

To ensure the global convergence of (7.24) towards zero for any u_k , not only the term PB has to vanish (input independence) but also the resulting dynamical equation must converge towards zero (global convergence).

PROPOSITION 2

The state reconstruction error Equation (7.24) can be made input independent whenever $\text{rank}(CB) = \text{rank}(B) = r$.

Indeed, according to the definition of P , the equality $PB = 0$ entails that Q must be subject to:

$$B = QCB \quad (7.24)$$

Proposition 2 ensures the existence of the solution Q of (7.25) and its general expression is:

$$Q = B(CB)^\dagger + Y(\mathbf{1}_m - (CB)(CB)^\dagger) \quad (7.25)$$

with Y an arbitrary matrix.

Then, whenever Q satisfies (7.26), $PB = 0$ and so (7.24) turns into an input independent dynamics:

$$\varepsilon_{k+1} = \sum_{i=1}^N \xi_k^i (PA_i - L_i C) \varepsilon_k \quad (7.26)$$

This is a polytopic formulation of the error of synchronization. On the basis of the same reasoning as in Section 7.3.3, the following theorem can thereby be stated:

THEOREM 4

IIGS of the message-embedded scheme (7.22) and (7.23) is achieved whenever the following conditions are satisfied:

1. $\text{rank}(CB) = \text{rank}(B) = r$
2. the set of LMIs:

$$\begin{bmatrix} P_i & (PA_i)^\top G_i^\top - C^\top F_i^\top \\ G_i(PA_i) - F_i C & G_i + G_i^\top - P_j \end{bmatrix} > 0 \quad (7.27)$$

is feasible for all $(i, j) \in \{1, \dots, N\} \times \{1, \dots, N\}$

The G_i 's, P_i 's, and F_i 's are unknown matrices of appropriate dimensions. The resulting gains L_i are given by $L_i = G_i^{-1}F_i$.

PROOF See Ref. [27]. Let us just mention that the rank Condition 1 ensures $(CB)^\dagger$ to exist in (7.26), and so $PB = 0$ to be satisfied. ■

7.4.4 Plaintext Recovering

The remaining problem lies in the plaintext m_k recovering. Proposition 2 provides the conditions and a way to recover the original plaintext m_k .

PROPOSITION 3

If (7.22) and (7.23) are IIGS, one has $\hat{d}_{k+1} = u_k$ and $\hat{m}_{k+1} = m_k$, where $\hat{d}_k$ and $\hat{m}_k$ obey the following recursions:

$$\hat{d}_{k+1} = (CB)^\dagger(y_{k+1} - CA(\rho_k)\hat{x}_k - CE(\rho'_k)) \quad (7.29a)$$

$$\hat{m}_{k+1} = e^{-1}(\hat{x}_k, \hat{d}_{k+1}) \quad (7.29b)$$

PROOF On one hand, if (7.22) and (7.23) are IIGS, then $\hat{x}_k = x_k$. Furthermore, premultiply the dynamical equation of (7.22) by C , then multiply the left-hand side by $(CB)^\dagger$ (whose existence is ensured by the rank Condition 1) of Theorem 4 yields:

$$u_k = (CB)^\dagger(y_{k+1} - CA(\rho_k)x_k - CE(\rho'_k)) \quad (7.30)$$

Identifying (7.30) and (7.29a) yields $\hat{d}_{k+1} = u_k$.

On the other hand, if (7.22) and (7.23) are IIGS, then the equality $\hat{x}_k = x_k$ still holds. Furthermore, $\hat{x}_k = x_k$ and $\hat{d}_{k+1} = u_k$ turn the equality $m_k = e^{-1}(x_k, u_k)$ (derived from the definition of e) into $m_k = e^{-1}(\hat{x}_k, \hat{d}_{k+1})$. Identifying this equality with (7.29b) yields $\hat{m}_{k+1} = m_k$. This completes the proof. ■

Let us note that $\hat{m}_{k+1} = m_k$ necessarily leads to a delay when attempting to recover m_k . It stems from the fact that (7.22) has a relative degree equalling one with respect to u_k .

7.5 Illustrative Examples

7.5.1 Unknown Input Observer Design

An experiment of a message-embedded chaos synchronization scheme involving a transmitter of the form (7.22) is conducted. The distinct matrices of the dynamics are:

$$A(\rho_k) = \begin{bmatrix} \rho_k & 1 & 0 \\ 1 & 0 & 1 \\ -0.7 & 0 & 0.01 \end{bmatrix}, \quad B = [0.1 \quad 1 \quad 1]^T, \\ C = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E = \mathbf{0}_{3 \times 1}$$

The nonlinear time-varying parameter ρ_k verifies $\rho_k = x_k^1 x_k^3, x_k^i$ being the i th component of the state vector x_k . Note that the special choice of the output matrix C causes ρ_k to be available at each discrete time k as required. The map exhibits a chaotic motion which entails that ρ_k is bounded. Simulation shows that ρ_k ranges between $\underline{\rho} = 0.5627$ and $\bar{\rho} = 0.8028$. Consequently, A lies in a compact set which may be embedded in a polytope and may be expressed in the form of (7.2). The corresponding vertices are given by:

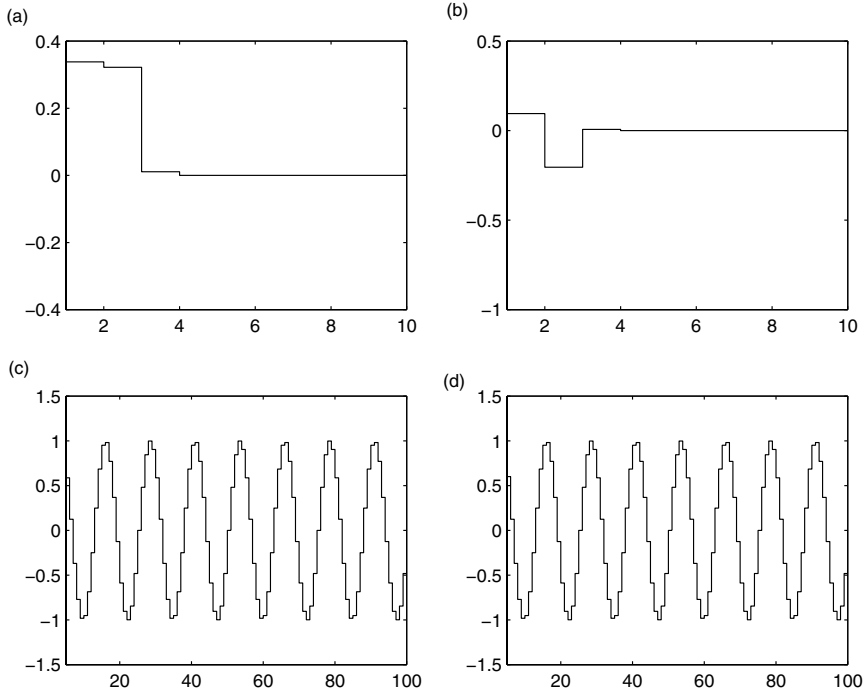
$$A_1 = \begin{bmatrix} 0.5627 & 1 & 0 \\ 1 & 0 & 1 \\ -0.7 & 0 & 0.01 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.8028 & 1 & 0 \\ 1 & 0 & 1 \\ -0.7 & 0 & 0.01 \end{bmatrix}$$

A one-dimensional information m_k (the plaintext) is embedded in the chaotic motion through u_k which results from an invertible coding function e . The plaintext consists of a sampled sine wave $m_k = M_m \sin(2\pi k \tilde{f})$, with $\tilde{f} = 0.08$.

To first recover u_k , the receiver is designed according to (7.23) and the Theorem 4. The gain Q is computed from (7.26) with a null arbitrary matrix Y . The gains L_i are computed from (7.28) using a standard LMI solver. Finally, the plaintext m_k is recovered by applying (7.29). Simulation results are presented in Figure 7.2 and highlight the consistence of the theoretical developments.

7.5.2 Real-Time Private Communication Experiment

A chaotic encryption experiment is conducted within a real-time video communication context. For that purpose, a well-known GNU licensed program for video transmission over the Internet, *VIC*, has been chosen.

**FIGURE 7.2**

(a) Error $\|x_k - \hat{x}_k\|$, (b) error $m_k - \hat{m}_k$, (c) plaintext m_k , and (d) recovered plaintext $\hat{m}_k$.

A video tape is connected to a computer through the S-Video port of a *Miro Studio PCTV RaveTM* card. Consequently, the video tape data are dynamically captured by the TV-card and VIC manages the transmission protocol over the Internet. The card has been configured to work under Linux and is driven by the bttv and video4linux drivers included in the VIC package. VIC's default package contains an implementation of DES cryptographic algorithm. However, it also allows to introduce new cryptographic schemes. The encrypting scheme corresponding to the message-embedded framework developed in Section 7.4 has been introduced. The well-known second-order Markov piecewise linear map has been considered as the chaos generator. It induces a chaotic behavior which is reckoned to get some good statistical properties for encryption purposes. Here, the key is the parameterization of the chaotic map. The design of the transmitter-receiver obeys the Equation (7.22), Equation (7.23), and Theorem 4.

With the same key in both sides, the video is correctly displayed in the receiver side. Using slightly different key parameters on both sides of the communication system, the image is badly decrypted, which highlights the sensitivity of the algorithm to key parameters. Two capture screens are shown in Figure 7.3.

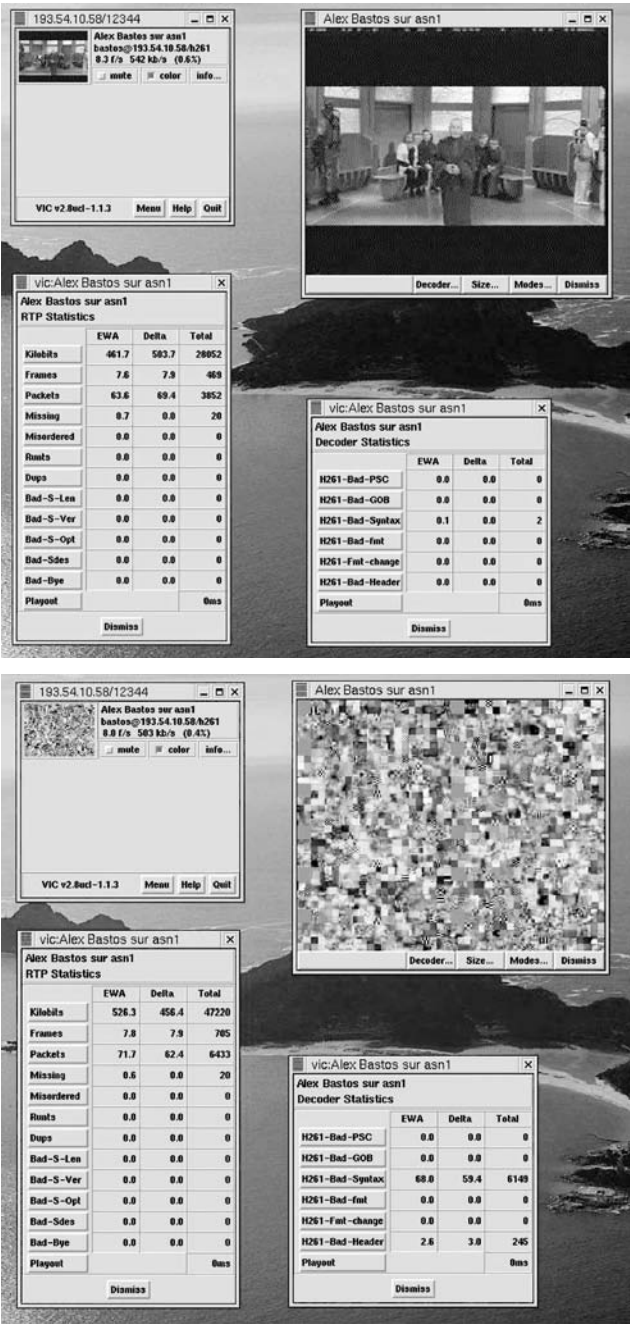


FIGURE 7.3
Decoder capture screens: matched and mismatched keys.

7.6 Conclusion

Chaos synchronization has been tackled by considering the problem as a special case of an observer design. The general class of considered dynamical systems includes, in particular, Lur'e systems, piecewise linear systems, and output injection with time-varying dynamical matrix. Both message-free and message-embedded synchronization issues have been treated. Owing to the chaotic motions-specific properties of keeping the underlying state variables in a compact domain, the dynamics can be written in a polytopic form. Hence, the synchronization can be carried out by designing, respectively, a polytopic observer or an unknown polytopic observer. Global synchronization is ensured by a special Liapunov approach. The Liapunov function is a PDLF called poly-quadratic, with a structure similar to that of the polytopic system description. The gains of the observer are computed by a systematic procedure involving the solutions of an LMI set.

References

1. G. Becker, Additional results on parameter-dependent controllers for lpv systems, in *Proceedings of the IFAC World Congress*, San Fransisco, 1996.
2. V.D. Blondel, E.D. Sontag, M. Vidyasagar, and J.C. Willems, Open problems in mathematical systems and control theory, *Communication and Control Engineering*. Springer Verlag, 1999.
3. I.I. Blekhman, A.L. Fradkov, H. Nijmeijer, and A.Y. Pogromsky, On self-synchronization and controlled synchronization, *Syst. Contr. Lett.*, 31 (5), 299–305, 1997.
4. S.-K. Chang, W.-T. You, and P.-L. Hsu, Design of general structured observers for linear systems with unknown inputs, *J. Franklin Inst.*, 334 (2), 213–232, 1997.
5. Chen, Ed., *Controlling Chaos and Bifurcations in Engineering Systems*, CRC Press, 1999.
6. J. Daafouz and J. Bernussou, Parameter dependent Lyapunov functions for discrete time systems with time varying parametric uncertainties, *Syst. Contr. Lett.*, 43, 355–359, 2001.
7. J. Daafouz and G. Millerioux, Poly-quadratic stability and global chaos synchronization of discrete time hybrid systems, *Special Issue of Mathematics and Computers in Simulation*, 58, 295–307, 2002.
8. J. Daafouz, G. Millerioux, and C. Iung, A poly-quadratic stability based approach for switched systems, *Int. J. Contr.*, 75, 1302–1310, 2002.
9. J. Daafouz, P. Riedinger, and C. Iung, Stability analysis and control synthesis for switched systems: a switched Lyapunov approach, *IEEE Trans. Automatic Contr.*, 2002.

10. M. Darouach, M. Zazadinski, and S.J. Xu, Full-order observers for linear systems with unknown inputs, *IEEE Trans. Automat. Contr.*, 39 (3), 606–609, 1994.
11. A. De Angeli, R. Genesio, and A. Tesi, Dead-beat chaos synchronization in discrete-time systems, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 42 (1), 54–56, 1995.
12. P. Gahinet, P. Apkarian, and M. Chilali, Affine parameter dependent Lyapunov functions and real parametric uncertainty, *IEEE Trans. Automat. Contr.*, 41, 436–442, 1996.
13. Y. Guan and M. Saif, A novel approach to the design of unknown input observers, *IEEE Trans. Automat. Contr.*, 36 (5), 632–635, 1991.
14. M. Hasler, Synchronization of chaotic systems and transmission of information, *Int. J. Bifurcat. Chaos*, 8 (4), 1998.
15. H.J.C. Huijberts and H. Nijmeijer, An observer view on synchronization, in A. Isidori, F. Lamnabhi-Laguarrigue, and W. Respondek, Eds., *Nonlinear Control in the year 2000*, Springer Verlag, 2000, pp. 509–520.
16. H.J.C. Huijberts, T. Lilge, and H. Nijmeijer, Nonlinear discrete-time synchronization via extended observers, *Int. J. Bifurcat. Chaos*, 11 (7), 1997–2006, 2001.
17. M. Itoh, C.W. Wu, and L.O. Chua, Communications systems via chaotic signals from a reconstruction viewpoint, *Int. J. Bifurcat. Chaos*, 7 (2), 275–286, 1997.
18. G. Jakimoski and L. Kocarev, Chaos and cryptography: block encryption ciphers based on chaotic maps, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 48 (2), 163–169, 2001.
19. L. Kocarev, Chaos-based cryptography: a brief overview, *IEEE Circ. Syst. Mag.*, 1 (3), 6–21, 2001.
20. S.H. Lee, J. Kong, and J.H. Seo, Observers for bilinear systems with unknown inputs and application to superheater temperature control, *Contr. Eng. Pract.*, 5 (4), 493–506, 1997.
21. T.-L. Liao and N.-S. Huang, An observer-based approach for chaotic synchronization with applications to secure communications, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 46 (9), 1144–1150, 1999.
22. K.-Y. Lian, T.-S. Chiang, and P. Liu, Discrete-time chaotic systems: applications in secure communications, *Int. J. Bifurcat. Chaos*, 10 (9), 2193–2206, 2000.
23. G. Millerioux and J. Daafouz, Global chaos synchronization and robust filtering in noisy context, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 48 (10), 1170–1176, 2001.
24. G. Millerioux and J. Daafouz, Polytopic observer for global synchronization of systems with output measurable nonlinearities, *Int. J. Bifurcat. Chaos*, 13 (3), 703–712, 2003.
25. G. Millerioux, F. Amstett, and G. Bloch, Considering the attractor structure of chaotic maps for observer-based synchronization problems, *Mathematics and Computers in Simulation*, 68 (1), 67–85, 2005.
26. G. Millerioux and J. Daafouz, An observer-based approach for input independent global chaos synchronization of discrete-time switched systems, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 50 (10), 1270–1279, 2003.

27. G. Millerioux and J. Daafouz, Unknown input observers for message-embedded chaos synchronization of discrete-time systems, *Int. J. Bifurcat. Chaos*, 14 (4), 1357–1368, 2004.
28. H. Nijmeijer and I.M.Y. Mareels, An observer looks at synchronization, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 44, 882–890, 1997.
29. M.J. Ogorzalek, Taming chaos — part I: synchronization, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 40 (10), 693–699, 1993.
30. L.M. Pecora and T.L. Carroll, Synchronization in chaotic systems, *Phys. Rev. Lett.*, 64, 821–824, 1990.
31. L.M. Pecora and T.L. Carroll, Driving systems with chaotic signals, *Phys. Rev. A*, 44 (8), 2374–2383, 1991.
32. A. Pogromsky and H. Nijmeijer, Observer-based robust synchronization of dynamical systems, *Int. J. Bifurcat. Chaos*, 8 (11), 2243–2254, 1998.
33. W.J. Rugh, Analytical framework for gain scheduling, *IEEE Contr. Syst. Mag.*, 11 (1), 74–84, 1991.
34. J.S. Shamma and M. Athans, Analysis of gain scheduled control for nonlinear plants, *IEEE Trans. Automat. Contr.*, 35, 898–907, 1990.
35. J.S. Shamma and M. Athans, Guaranteed properties of gain scheduled control for linear parameter varying plants, *Automatica*, 27, 559–564, 1991.
36. H. Sira Ramirez, C.A. Ibanez, and Suarez-castanon, Exact state reconstructors in the recovery of messages encrypted by the states of nonlinear discrete-time chaotic systems, *Int. J. Bifurcat. Chaos*, 12 (1), 169–177, 2002.
37. H. Sira Ramirez and C. Cruz Hernandez, Synchronization of chaotic systems: a generalized hamiltonian approach, *Int. J. Bifurcat. Chaos*, 11 (5), 1381–1395, 2001.
38. Special Issue, Control of chaos and synchronization, *Syst. Contr. Lett.*, 31, 259–322, 1997.
39. Special Issue, Chaos synchronization and control: theory and applications, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 40, 853–1039, 1997.
40. Special Issue, Control and synchronization of chaos, *Int. J. Bifurcat. Chaos*, 10 (4), 2000.
41. Special Issue, Control of oscillations and chaos, *Special Issue Math. Comput. Simulat.*, 58, 283–505, 2002.
42. Takens F, Detecting strange attractors in fluid turbulence, in D. Rand and L.S. Young, Eds., *Dynamical Systems and Turbulences*, Springer-Verlag, Berlin, 1981.
43. F. Wu, An unified framework for lpv system analysis and control synthesis, in *Proceedings of the IEEE Conference on Decision and Control*, Sydney, Australia, December, 2000.
44. T. Yang, A survey of chaotic secure communication systems, *Int. J. Comput. Cogn.*, 2004 (available at <http://www.YangSky.com/yangijcc.htm>).
45. F. Yang and R.W. Wilde, Observers for linear systems with unknown inputs, *IEEE Trans. Automat. Contr.*, 33 (7), 677–681, 1988.
46. H. Zhou and X.-T. Ling, Problems with the chaotic inverse system encryption approach, *IEEE Trans. Circ. Syst. I: Fundam. Theor. Appl.*, 44 (3), 268–271, 1997.

8

Normal Forms of Nonlinear Control Systems

W. Kang and A. J. Krener

CONTENTS

8.1	Introduction	345
8.2	Linearly Controllable Systems	348
8.3	Linearly Uncontrollable Systems	356
8.4	Examples of Normal Form	367
8.4.1	The Normal Form of Ball and Beam	367
8.4.2	Engine Compressor	370
8.4.3	Controlled Lorenz Equation	373
8.5	Conclusions	375
	References	375

8.1 Introduction

Numerous papers were published during the last decade on the normal forms of nonlinear control systems with applications in bifurcation and its control. The approach is motivated by Poincaré's theory of normal forms [1] for classical dynamical systems using homogeneous transformations. In this paper, we summarize a variety of control system normal forms published in the literature so that the normal forms are derived in a similar framework with consistent notations. Before we get into technical details, the rest of this section reviews existing results on some related topics.

It is well known that there are several normal forms for a linear control system. If the system is controllable then the system can be transformed into controllable or controller normal form. If the system has a linear output map and is observable, then it can be transformed into an observable or

observer form. The nonlinear generalization of the linear controller normal forms were extensively studied during the 1980s, for instance, Krener [22], Hunt–Su [11], Jackubczyk–Respondek [10], and Brockett [3], etc. If a nonlinear control system admits a controller normal form, it can be transformed into a linear system by a change of coordinates and feedback. Therefore, the design of a locally stabilizing state feedback control law is a straightforward task. In such a case, we say the system is feedback linearizable. On the other hand, most nonlinear systems do not admit a controller normal form under change of coordinates and invertible state feedback. For systems that are not linearizable, the quadratic approximate version of controller normal form was introduced and discussed in Krener [21] and Krener–Karahan–Hubbard–Frezza [23]. It was proved that, for certain kinds of nonlinear systems, there exist a quadratic change of coordinates and quadratic feedback that transform the system into the linear approximation of the plant dynamics which is accurate to at least the second degree. In this case, we call the system quadratically equivalent to a linear system or quadratically feedback linearizable. However, most nonlinear systems do not admit such a linear approximation. Another way of linearizing a nonlinear control system is dynamic feedback linearization. Some nonlinear systems with more than one input can be linearized by a dynamic feedback even if they are not linearizable by a state feedback. However, it was proved that a dynamic feedback cannot completely linearize a nonlinear system with single input if it is not linearizable by a state feedback (see Charlet–Lévine–Marino [4]).

Until late 1980s, the problem of normal forms for nonlinear control systems that are not feedback linearizable was still largely open. On the other hand, the Poincaré normal form of nonlinear dynamic systems has been a successful theory with applications in the study of bifurcations and stability. Although the normal form of Poincaré was not applied to control systems, in Kang [12], the idea of Poincaré was applied to nonlinear control systems with a single input. A normal form was derived for the family of linearly controllable systems with a single input, including systems that are not feedback linearizable. In addition, it was proved in Kang [12] that a dynamic feedback is able to approximately linearize a controllable system to an arbitrary degree. Invariants were found in Kang [12] that uniquely determine the normal form of a control system. The homogeneous parts of degree d from two systems are equivalent under homogeneous transformations if and only if they have the same invariants. Part of the dissertation [12] were published in Kang–Krener [13] and Kang [14, 16].

Starting from early 1990s, the research on normal forms moved in several related but different directions. One active research direction is to find the normal forms of systems with uncontrollable linearizations. Several authors have made contributions to this subject. Quadratic normal forms of systems with uncontrollable linearization were developed by Kang [15, 17, 18]. The results were generalized to higher degree terms

by Fitch [6] and Tall–Respondek [28, 31] for systems affine in control. In Krener–Kang–Cheng [25], the normal form and invariants of nonlinear control systems with a single input, not necessarily affine in control, is achieved through the third degree. In the following sections, the result is generalized to homogeneous terms of arbitrary degree. The proof in Reference [25] is constructive, which is different from the existence proof adopted in most previous published works. The same constructive proof is adopted in this chapter and generalized to higher degrees. Similar to Poincaré’s theory, the normal form of a control system is invariant under homogeneous transformations of the same degree. However, a normal form of degree k is not unique under transformations of degree less than k . If a normal form is unique under transformation of arbitrary degree, it is called a canonical form. Tall–Respondek [34] solved the problem of canonical form for single-input and linearly controllable systems.

For multi-input systems, their nonlinear normal forms and invariants were first studied in Kang [12]. The quadratic normal form and quadratic invariants were derived in Reference [12] for linearly controllable systems in which the controllability indices equal each other. Without any assumption on the controllability indices, Tall–Respondek [32] found a normal form of arbitrary degree for linearly controllable systems with two inputs. The results were further generalized by Tall [33] for linearly controllable systems with any number of inputs. The normal form was derived for homogeneous parts of arbitrary degree.

Barbot, Monaco, and Normand-Cyrot [2] derived a linear and quadratic normal form for linearly controllable discrete-time systems. Quadratic and cubic normal forms were derived by Krener–Li [24] for general discrete-time systems of both linearly controllable and uncontrollable systems. The approaches adopted by these two groups [2, 24] are different. As a result, the normal forms derived in the two papers are different for linearly controllable systems.

The application of normal forms and invariants of control systems is another active research topic. On the basis of normal forms, bifurcations and its control were studied by several authors. In Kang [17–19], bifurcations and their classification for both open-loop and closed-loop systems were studied for systems with a single uncontrollable mode. In Krener–Kang–Cheng [25], control bifurcation for parameterized state feedback was studied. Hamzi–Kang–Barbot [8] used normal forms and invariants to characterize the orientation and stability of periodic trajectories in a Hopf bifurcation under state feedback. Controllability and accessibility of normal forms were addressed in Kang–Xiao–Tall [21]. Bifurcations and their control for discrete-time systems is addressed in References [7, 24].

As an application of canonical form, Respondek–Tall [28, 29] studied the symmetry of nonlinear systems. For linearly controllable and analytical systems that are not feedback linearizable, the group of stationary

symmetries contains at most two elements and the group of nonstationary symmetries consist of at most two 1-parameter families. This surprising result follows from the canonical form obtained for single-input systems by Tall–Respondek [34]. Respondek [30] establishes the relationship between flatness and symmetries for two classes of systems: feedback linearizable systems and systems equivalent to the canonical contact system for curves. For these two classes of systems, the minimal flat outputs determine local symmetries and vice versa.

8.2 Linearly Controllable Systems

In this section, a control system with a scalar input is defined by the following equation,

$$\dot{x} = f(x, u) \quad (8.1)$$

where $x \in \mathbb{R}^n$ is the state variable and $u \in \mathbb{R}$ is the control input. Occasionally, it is notationally convenient to denote the control input u by x_{n+1} . We assume that the function $f(x, u)$ is C^k for sufficiently large k . An equilibrium is a pair (x_e, u_e) that satisfies

$$f(x_e, u_e) = 0 \quad (8.2)$$

An equilibrium state x_e is one for which there exists an u_e so that (x_e, u_e) is an equilibrium. Consider the linearization of (8.1) at (x_e, u_e) ,

$$\delta \dot{x} = F \delta x + G \delta u, \quad F = \frac{\partial f}{\partial x}(x_e, u_e), \quad G = \frac{\partial f}{\partial u}(x_e, u_e). \quad (8.3)$$

A control system (8.1) is *linearly controllable* at (x_e, u_e) if its linearization (8.3) is controllable. The linear system (8.3) is controllable if

$$\text{rank} \begin{bmatrix} G & FG & F^2G & \cdots & F^{n-1}G \end{bmatrix} = n$$

In this section, the focus is on the normal form of linearly controllable systems. The normal form of a system with an uncontrollable linearization is addressed in Section 8.3. By a translation of the (x, u) coordinate system, we can assume that the equilibrium (x_e, u_e) is the origin $(0, 0)$.

Following the method of Poincaré, we derive the normal form of (8.1) by applying homogeneous transformations to the following Taylor expansion of (8.1)

$$\dot{x} = Fx + Gu + \sum_{k=2}^d f^{[k]}(x, u) + O(x, u)^{d+1} \quad (8.4)$$

In (8.4), $f_i^{[k]}(x, u)$ is a vector field in $\mathbb{R}^n$ whose components are homogeneous polynomials of degree k in (x, u) . For each homogeneous part, we apply homogeneous transformations to derive the normal form. For control systems, the transformation group includes both changes of state coordinates and invertible state feedbacks. A linear transformation is defined by

$$z = Tx, \quad v = Kx + Lu \quad (8.5)$$

where $T \in \mathbb{R}^{n \times n}$ is an invertible matrix, $K \in \mathbb{R}^n$ is a row vector, and $L \neq 0$ is a scalar. A transformation of degree $k > 1$ is defined by

$$z = x - \phi^{[k]}(x), \quad v = u - \alpha^{[k]}(x, u) \quad (8.6)$$

A transformation of degree k may change the homogeneous term $f^{[d]}(x, u)$ in (8.4) for $d \geq k$. However, a transformation (8.6) does not change any term of degree less than k . Similar to the derivation of Poincaré normal form, we derive the linear normal form of an equilibrium of a control system using a linear transformation. Then a quadratic transformation is used to derive the quadratic normal form. Because the quadratic transformation leaves the linear part invariant, the derivation of quadratic normal form does not change the linear normal form. In general, if the normal forms of $f^{[1]}(x, u), \dots, f^{[k-1]}(x, u)$ have been derived, a transformation of degree k is used to derive the normal form for $f^{[k]}$ in (8.4), which leaves the normal form of $f^{[l]}(x, u)$ invariant for $1 \leq l \leq k-1$.

It is well known that by linear transformation (8.6), a linear control system

$$\dot{x} = Fx + Gu \quad (8.7)$$

can be brought to the *Brunovsky form*

$$\dot{z} = Az + Bv \quad (8.8)$$

where A and B are of the form

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{n \times n}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{n \times 1} \quad (8.9)$$

The existence of such a linear transformation is proved in many textbooks of linear control systems.

Consider a linearly controllable system (8.4). We adopt the Brunovsky form as the linear normal form. There exists a linear transformation that brings (8.4) to the form

$$\dot{x} = Ax + Bu + f^{[2]}(x, u) + O(x, u)^3 \quad (8.10)$$

where (A, B) are defined by (8.9). In the following, we use a quadratic transformation

$$z = x - \phi^{[2]}(x), \quad v = u - \alpha^{[2]}(x, u) \quad (8.11)$$

to simplify the quadratic nonlinear part of the system. There are two basic operations, *pull up* and *push down*, which are used to achieve this.

Consider a part of the dynamics

$$\begin{aligned} \dot{x}_{i-1} &= x_i + \cdots \\ \dot{x}_i &= x_{i+1} + cx_j x_k + \cdots \end{aligned} \quad (8.12)$$

where $2 \leq i \leq n$, $1 \leq j \leq k \leq n+1$, recall $x_{n+1} = u$. The $+\cdots$ indicates other quadratic and higher degree terms. The other quadratic terms will not be changed by the operations that we perform. The higher terms may be changed but we are not interested in them at this time.

If $j < k-1$ we can *pull up* the quadratic term by defining

$$\begin{aligned} z_i &= x_i - cx_j x_{k-1} \\ z_l &= x_l \quad \text{if } l \neq i \end{aligned} \quad (8.13)$$

Its inverse transform satisfies

$$x_i = z_i + cz_j z_{k-1} + O(z)^3 \quad (8.14)$$

Then the dynamics becomes

$$\begin{aligned} \dot{z}_{i-1} &= z_i + cz_j z_{k-1} + \cdots \\ \dot{z}_i &= z_{i+1} - cz_{j+1} z_{k-1} + \cdots \end{aligned} \quad (8.15)$$

and all the other quadratic terms remain the same. In each of the new quadratic terms, the two indices are closer together than the two indices of

the original quadratic term. If $j = k - 1$, we can pull up the quadratic term by defining

$$\begin{aligned} z_i &= x_i - \frac{c}{2}x_jx_k \\ z_l &= x_l \quad \text{if } l \neq i \end{aligned} \quad (8.16)$$

then the dynamics becomes

$$\begin{aligned} \dot{z}_{i-1} &= z_i + \frac{c}{2}z_jz_k + \cdots \\ \dot{z}_i &= z_{i+1} + \cdots \end{aligned} \quad (8.17)$$

and all the other quadratic terms remain the same. The two indices of the new quadratic term are identical.

Note also that in either case if $i = 1$, we can still pull up and there is no z_{i-1} dynamics to be concerned with, so a term disappears.

By pulling up all the quadratic terms until the two indices are equal, we obtain

$$\dot{x}_i = x_{i+1} + \sum_{j=1}^{n+1} \epsilon_{ij} x_j^2 + \cdots \quad (8.18)$$

where x denotes the new state coordinate after the pull up process. This form can be simplified further by the other operation on the dynamics, *push down*. Consider a piece of the dynamics

$$\begin{aligned} \dot{x}_i &= x_{i+1} + cx_jx_k + \cdots \\ \dot{x}_{i+1} &= x_{i+2} + \cdots \end{aligned} \quad (8.19)$$

where $1 \leq i \leq n - 1$ and $1 \leq j \leq k \leq n$. Define

$$\begin{aligned} z_{i+1} &= x_{i+1} + cx_jx_k \\ z_l &= x_l \quad \text{if } l \neq i + 1 \end{aligned} \quad (8.20)$$

Its inverse transformation satisfies

$$x_{i+1} = z_{i+1} - cz_jz_k + O(z)^3 \quad (8.21)$$

The transformation (8.20) yields

$$\begin{aligned} \dot{z}_i &= z_{i+1} + \cdots \\ \dot{z}_{i+1} &= z_{i+2} + cz_{j+1}z_k + cz_jz_{k+1} + \cdots \end{aligned} \quad (8.22)$$

and all the other quadratic terms remain unchanged. Notice that if $i + 1 = n$ then we can absorb any quadratic term into the control using feedback. The terms in (8.19), where $1 \leq j \leq k \leq i + 1$, can be pushed down repeatedly and absorbed in the control.

If the control appears in the derivative of one of the states, then we cannot push that term down any further since the control need not be differentiable. So, if the term cx_jx_k appears in the equation for $\dot{z}_i$ with j or k greater than $i + 1$ and we try to repeatedly push it down then the control will appear before we reach the equation for $\dot{z}_n$. For this reason, we only push down a quadratic term x_jx_k with both j and k less than or equal to $i + 1$. As a result, the system (8.18) is transformed into the following quadratic normal form.

$$\begin{aligned}\dot{x}_i &= x_{i+1} + \sum_{j=i+2}^{n+1} \epsilon_{i,j} x_j x_i + O(x, u)^3, \quad \text{for } 1 \leq i \leq n-1 \\ \dot{x}_n &= u + O(x, u)^3\end{aligned}\tag{8.23}$$

where x represents the new state coordinates after the push down process.

Example 1

The following is the quadratic normal form of the general two-dimensional linearly controllable system.

$$\begin{aligned}\dot{x}_1 &= x_2 + \epsilon_{1,3} x_2^2 + O(x, u)^3 \\ \dot{x}_2 &= u + O(x, u)^3\end{aligned}\tag{8.24}$$

Notice there is only one coefficient that cannot be normalized to zero and this is the invariant of the system under quadratic transformations.

The following is the quadratic normal form of the general three-dimensional linearly controllable system:

$$\begin{aligned}\dot{x}_1 &= x_2 + \epsilon_{1,3} x_3^2 + \epsilon_{1,4} u^2 + O(x, u)^3 \\ \dot{x}_2 &= x_3 + \epsilon_{2,4} u^2 + O(x, u)^3 \\ \dot{x}_3 &= u + O(x, u)^3\end{aligned}\tag{8.25}$$

Now there are three coefficients that cannot be normalized to zero and these are the invariants of the system under quadratic transformations.

For the rest of the section, we use pull up and push down to prove the following theorem on general normal forms.

THEOREM 1

Suppose (8.1) is linearly controllable. Suppose the vector field $f(x, u)$ is C^{d+1} . Then by a change of coordinates and a state feedback, (8.1) can be transformed into the following normal form

$$\begin{aligned}\dot{z} &= Az + Bv + \sum_{k=2}^d \tilde{f}^{[k]}(z) + O(z, v)^{d+1} \\ \tilde{f}_i^{[k]}(z) &= \sum_{j=i+2}^{n+1} \epsilon_{i,j}^{[k-2]}(\bar{z}_j) z_j^2\end{aligned}\tag{8.26}$$

where (A, B) is in the Brunovsky form. The coefficient $\epsilon_{i,j}^{[k-2]}(\bar{z}_j)$ is a homogeneous polynomial of degree $k-2$ in the variable $\bar{z}_j = (z_1, z_2, \dots, z_j)$. When there are no terms in the sum then it is zero as in

$$\tilde{f}_n^{[k]}(z) = \sum_{j=n+2}^{n+1} \epsilon_{i,j}^{[k-2]}(\bar{z}_j) z_j^2 = 0\tag{8.27}$$

PROOF Consider the expansion (8.4). The proof follows by mathematical induction. We have derived the linear and quadratic normal forms. Suppose that all homogeneous parts of degree less than m in (8.4) are transformed into their normal forms, consider the homogeneous part $f^{[m]}(x)$ in (8.4). A part of the dynamics has the form

$$\begin{aligned}\dot{x}_{i-1} &= x_i + \sum_{k=2}^{m-1} \tilde{f}_{i-1}^{[k]}(x, u) + \dots \\ \dot{x}_i &= x_{i+1} + \sum_{k=2}^{m-1} \tilde{f}_i^{[k]}(x, u) + cx_{j_1}x_{j_2} \dots x_{j_m} + \dots\end{aligned}\tag{8.28}$$

where $2 \leq i \leq n$, $1 \leq j_1 \leq j_2 \leq \dots \leq j_m \leq n+1$, recall $x_{n+1} = u$. The $+\dots$ stands for other homogeneous terms of degree m and higher. The other terms of degree m will not be affected by the operations that we perform and we ignore the higher degree terms. A transformation of degree m does not change the normal form of degree less than m .

If $j_{m-1} < j_m - 1$ we can pull up the degree m term by defining

$$\begin{aligned}z_i &= x_i - cx_{j_1}x_{j_2} \dots x_{j_{m-1}}x_{j_m-1} \\ z_l &= x_l, \quad \text{for } l \neq i\end{aligned}\tag{8.29}$$

then the dynamics becomes

$$\begin{aligned}
 \dot{z}_{i-1} &= z_i + \sum_{k=2}^{m-1} \tilde{f}_{i-1}^{[k]}(z, u) + cz_{j_1}z_{j_2} \cdots z_{j_{m-1}}z_{j_m-1} + \cdots \\
 \dot{z}_i &= z_{i+1} + \sum_{k=2}^{m-1} \tilde{f}_i^{[k]}(z, u) - cz_{j_1+1}z_{j_2} \cdots z_{j_{m-1}}z_{j_m-1} \\
 &\quad - cz_{j_1}z_{j_2+1} \cdots z_{j_{m-1}}z_{j_m-1} - \cdots - cz_{j_1}z_{j_2} \cdots z_{j_{m-1}+1}z_{j_m-1} + \cdots \\
 &= z_{i+1} + \sum_{k=2}^{m-1} \tilde{f}_i^{[k]}(z, u) - c \sum_{k=1}^{m-1} \frac{z_{j_1}z_{j_2} \cdots z_{j_{m-1}}z_{j_m-1}}{z_{j_k}} z_{j_{k+1}} + \cdots
 \end{aligned} \tag{8.30}$$

and all the other degree m terms remain the same. Notice that the two largest indices of the new degree m terms are closer together than those of the original degree m term.

If $j_{m-p-1} < j_{m-p} = j_{m-p+1} = \cdots j_{m-1} = j_m - 1$, we can pull up the degree m term by defining

$$\begin{aligned}
 z_i &= x_i - \frac{c}{p+1} x_{j_1} x_{j_2} \cdots x_{j_{m-p-1}} x_{j_m-1}^{p+1} \\
 z_l &= x_l, \quad \text{for } l \neq i
 \end{aligned} \tag{8.31}$$

then the dynamics becomes

$$\begin{aligned}
 \dot{z}_{i-1} &= z_i + \sum_{k=2}^{m-1} \tilde{f}_{i-1}^{[k]}(z, u) + \frac{c}{p+1} z_{j_1}z_{j_2} \cdots z_{j_{m-p-1}}z_{j_m-1}^{p+1} + \cdots \\
 \dot{z}_i &= z_{i+1} + \sum_{k=2}^{m-1} \tilde{f}_i^{[k]}(z, u) - \frac{c}{p+1} \sum_{k=1}^{m-p-1} \frac{z_{j_1}z_{j_2} \cdots z_{j_{m-p-1}}z_{j_m-1}^{p+1}}{z_{j_k}} z_{j_{k+1}} + \cdots
 \end{aligned} \tag{8.32}$$

and all the other degree m terms remain the same. Notice that the two largest indices of the new degree m terms are identical.

In any case, if $i = 1$ then we can still pull up and there is no z_{i-1} dynamics to be concerned with, so a term disappears.

By pulling up all the degree m terms until their two largest indices are identical, we obtain

$$\dot{x}_i = x_{i+1} + \sum_{k=2}^{m-1} \tilde{f}_i^{[k]}(x, u) + \sum_{j=1}^{n+1} \epsilon_{ij}^{[m-2]}(\bar{x}_j) x_j^2 + \cdots \tag{8.33}$$

which is almost the normal form (8.26).

By pushing down, we can make $\epsilon_i^j = 0$ for $1 \leq j \leq i + 1$. Consider a piece of the dynamics,

$$\begin{aligned}\dot{x}_i &= x_{i+1} + \sum_{k=2}^{m-1} \tilde{f}_i^{[k]}(x, u) + cx_{j_1}x_{j_2} \cdots x_{j_m} + \cdots \\ \dot{x}_{i+1} &= x_{i+2} + \sum_{k=2}^{m-1} \tilde{f}_{i+1}^{[k]}(x, u) + \cdots\end{aligned}\tag{8.34}$$

If $1 \leq j_1 \leq j_2 \leq \cdots \leq j_m \leq n$, define

$$\begin{aligned}z_{i+1} &= x_{i+1} + cx_{j_1}x_{j_2} \cdots x_{j_m} \\ z_l &= x_l, \quad \text{for } l \neq i + 1\end{aligned}\tag{8.35}$$

yielding

$$\begin{aligned}\dot{z}_i &= z_{i+1} + \sum_{k=2}^{m-1} \tilde{f}_i^{[k]}(z, u) + \cdots \\ \dot{z}_{i+1} &= z_{i+2} + \sum_{k=2}^{m-1} \tilde{f}_{i+1}^{[k]}(z, u) + c \sum_{k=1}^m \frac{z_{j_1}z_{j_2} \cdots z_{j_m}}{z_{j_k}} z_{j_k+1} + \cdots\end{aligned}\tag{8.36}$$

and all the other degree m terms remain unchanged. Notice that if $i + 1 = n$, then we can absorb the degree m terms into the control using feedback. The terms in (8.33) where $1 \leq j_1 \leq j_2 \leq \cdots \leq j_m \leq i + 1$ can be repeatedly pushed down and absorbed in the control. The result is (8.26). ■

Example 2

The following is the normal form up to the fourth degree for a general three-dimensional system:

$$\begin{aligned}\dot{x}_1 &= x_2 + \epsilon_{1,3}(x)x_3^2 + \epsilon_{1,4}(x, u)u^2 + O(x, u)^4 \\ \dot{x}_2 &= x_3 + \epsilon_{2,4}(x, u)u^2 + O(x, u)^4 \\ \dot{x}_3 &= u + O(x, u)^4\end{aligned}\tag{8.37}$$

where $\epsilon_{i,j} = \epsilon_{i,j}^{[0]} + \epsilon_{i,j}^{[1]} + \epsilon_{i,j}^{[2]}$ and $\epsilon_{i,j}^{[k]}$ is a homogeneous polynomial of degree k .

8.3 Linearly Uncontrollable Systems

In this section, we generalize the results of Section 8.2 to systems with uncontrollable linearization. Consider a control system (8.1). Assume that the controllability matrix of its linearization (8.3) has a rank $n_1 < n$. It is well known that by linear change of state coordinates and linear state feedback, the system can be brought to the form

$$\begin{bmatrix} \dot{x}_0 \\ \dot{x}_1 \end{bmatrix} = \begin{bmatrix} A_0 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ B_1 \end{bmatrix} u \\ + \sum_{k=2}^d \begin{bmatrix} f_0^{[k]}(x_0, x_1, u) \\ f_1^{[k]}(x_0, x_1, u) \end{bmatrix} + O(x_0, x_1, u)^{d+1}$$

where x_0 and x_1 are n_0 - and n_1 -dimensional, respectively; $n_0 + n_1 = n$, $u \in \mathbb{R}$, A_0 is in the block diagonal Jordan form, A_1, B_1 are in the Brunovsky form, and $f_r^{[d]}(x_0, x_1, u)$ is a vector field which is a homogeneous polynomial of degree d in its arguments. The linear change of coordinates that brings A_0 to the Jordan form may be complex, in which case some of the coordinates $x_{0,i}$ are complex. The complex coordinates come in conjugate pairs. The corresponding $f_{0,i}^{[k]}$ are complex valued and come in conjugate pairs. In some formulae, the control input is treated as a state variable $u = x_{0,n_1+1}$. A nonlinear vector field $f_r^{[k]}(x_0, x_1, u)$, $r = 0, 1$, has the following decomposition

$$f_r^{[k]}(x_0, x_1, u) = \sum_{|l|=k} f_r^{[l]}(x_0; x_1, u) \quad (8.38)$$

where $[l] = [l_0; l_1]$ is a multi-index and $f_r^{[l]}(x_0; x_1, u)$ denotes a polynomial vector field homogeneous of degree l_0 in x_0 and homogeneous of degree l_1 in (x_1, u) , $|l| = l_0 + l_1$. A homogeneous transformation of degree k has the following form

$$\begin{bmatrix} z_0 \\ z_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} - \begin{bmatrix} \phi_0^{[k]}(x_0, x_1) \\ \phi_1^{[k]}(x_0, x_1) \end{bmatrix} \quad (8.39)$$

$$v = u - \alpha^{[k]}(x_0; x_1, u)$$

We can expand it as follows:

$$\begin{bmatrix} z_0 \\ z_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} - \sum_{|l|=k} \begin{bmatrix} \phi_0^{[l]}(x_0; x_1) \\ \phi_1^{[l]}(x_0; x_1) \end{bmatrix} \quad (8.40)$$

$$v = u - \sum_{|l|=k} \alpha^{[l]}(x_0; x_1, u)$$

where $\phi_r^{[l]}(x_0; x_1)$ denotes a vector field that is homogeneous of degree l_0 in x_0 and homogeneous of degree l_1 in x_1 . Similarly, $\alpha^{[l]}(x_0; x_1, u)$ is a polynomial homogeneous of degree l_0 in x_0 and homogeneous of degree l_1 in (x_1, u) . Under a transformation (8.40), the degree $[l]$ terms are transformed into

$$\begin{aligned} \tilde{f}_0^{[l]}(z_0; z_1, v) &= f_0^{[l]}(z_0; z_1, v) - \frac{\partial \phi_0^{[l]}}{\partial z_0}(z_0; z_1) A_0 z_0 \\ &\quad - \frac{\partial \phi_0^{[l]}}{\partial z_1}(z_0; z_1) (A_1 z_1 + B_1 v_1) + A_0 \phi_0^{[l]}(z_0; z_1) \\ \tilde{f}_1^{[l]}(z_0; z_1, v) &= f_1^{[l]}(z_0; z_1, v) - \frac{\partial \phi_1^{[l]}}{\partial z_0}(z_0; z_1) A_0 z_0 \\ &\quad - \frac{\partial \phi_1^{[l]}}{\partial z_1}(z_0; z_1) (A_1 z_1 + B_1 v_1) \\ &\quad + A_1 \phi_1^{[l]}(z_0; z_1) + B_1 \alpha^{[l]}(z_0; z_1, v) \end{aligned} \quad (8.41)$$

This is still a homogeneous vector of degree $[l]$. We have proved the following lemma.

LEMMA 1

After the transformation (8.40), the new homogeneous part $\tilde{f}_0^{[l]}$ is completely determined by $f_0^{[l]}$ and $\phi_0^{[l]}(x_0; x_1)$. The new homogeneous part $\tilde{f}_1^{[l]}$ is completely determined by $f_1^{[l]}$, $\phi_1^{[l]}(x_0; x_1)$, and $\alpha^{[l]}(x_0; x_1, v)$.

According to the lemma, each component of the term, $\tilde{f}_r^{[l]}$, that is homogeneous of degree $[l]$ can be considered separately in the derivation of the normal form. Following Poincaré, (8.41) is called a homological equation. In the derivation of the normal form, the quadratic transformation is first applied to (8.38) to derive the normal form of $f^{[2]}(x_0, x_1, u)$. Then, a cubic transformation is used to derive the normal form of the cubic part. In general, after the normal form of degree less than k has been found, a homogeneous transformation of degree k is used to derive the normal form of $f^{[k]}(x_0, x_1, u)$.

THEOREM 2

Consider a control system (8.38). There exist homogeneous transformations of the form (8.39) with $k = 2, 3, \dots, d$ that transform the system (8.38) into the normal form

$$\begin{bmatrix} \dot{z}_0 \\ \dot{z}_1 \end{bmatrix} = \begin{bmatrix} A_0 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \end{bmatrix} + \begin{bmatrix} 0 \\ B_1 \end{bmatrix} v + \sum_{k=2}^d \begin{bmatrix} \tilde{f}_0^{[k]}(z_0, z_1, v) \\ \tilde{f}_1^{[k]}(z_0, z_1, v) \end{bmatrix} + O(z_0, z_1, v)^{d+1} \quad (8.42)$$

where $\tilde{f}_0^{[k]}, \tilde{f}_1^{[k]}$ have the following decomposition:

$$\begin{aligned} \tilde{f}_0^{[k]}(z_0, z_1, v) &= \tilde{f}_0^{[k;0]}(z_0) + \tilde{f}_0^{[k-1;1]}(z_0; z_{1,1}) + \sum_{l_1=2}^k \tilde{f}_0^{[k-l_1;l_1]}(z_0; z_1, v) \\ \tilde{f}_1^{[k]}(z_0, z_1, v) &= \sum_{l_1=2}^k \tilde{f}_1^{[k-l_1;l_1]}(z_0; z_1, v) \end{aligned} \quad (8.43)$$

The vector field $\tilde{f}_0^{[k;0]}(z_0)$ is in the Poincaré normal form

$$\tilde{f}_{0,i}^{[k;0]}(z_0) = \sum_{\substack{|j|=k \\ j \cdot \lambda = \lambda_i}} \beta_{i,j} z_0^j \quad (8.44)$$

where $j = (j_1, \dots, j_{n_0})$ is a multi-index of nonnegative integers, $|j| = j_1 + \dots + j_{n_0}$, $j \cdot \lambda = j_1 \lambda_1 + \dots + j_{n_0} \lambda_{n_0}$, and $z_0^j = z_{0,1}^{j_1} \dots z_{0,n_0}^{j_{n_0}}$. The other vector fields are as follows:

$$\begin{aligned} \tilde{f}_{0,i}^{[k-1;1]}(z_0; z_{1,1}) &= \gamma_i^{[k-1]}(z_0) z_{1,1} \quad i = 1, \dots, n_0 \\ \tilde{f}_{0,i}^{[k-l_1;l_1]}(z_0; z_1, v) &= \sum_{j=1}^{n_1+1} \delta_{i,j}^{[k-l_1;l_1-2]}(z_0; \bar{z}_{1,j}) z_{1,j}^2 \quad i = 1, \dots, n_0 \quad l_1 = 2, \dots, k \\ \tilde{f}_{1,i}^{[k-l_1;l_1]}(z_0; z_1, v) &= \sum_{j=i+2}^{n_1+1} \epsilon_{i,j}^{[k-l_1;l_1-2]}(z_0; \bar{z}_{1,j}) z_{1,j}^2 \quad i = 1, \dots, n_1 \end{aligned} \quad (8.45)$$

where j is a scalar index, $z_{1,n_1+1} = v$, $\bar{z}_{1,j} = (z_{1,1}, z_{1,2}, \dots, z_{1,j})$ and $\delta_{i,j}^{[k-l_1;l_1-2]}(z_0; \bar{z}_{1,j})$, $\epsilon_{i,j}^{[k-l_1;l_1-2]}(z_0; \bar{z}_{1,j})$ are polynomials homogeneous of degree $k - l_1$ in z_0 and homogeneous of degree $l_1 - 2$ in (z_1, v) .

PROOF Suppose the homogeneous vector fields $f^{[k]}(x_0, x_1, u)$, for all $k \leq d-1$, are already in normal form. Consider the homogeneous term of degree d . A transformation of degree d does not change the homogeneous parts of degree less than d . It changes the terms of degree greater or equal to d . Because of Lemma 1, we can derive the normal form for each homogeneous part $f_r^{[l]}$ separately.

Consider $f_1^{[d;0]}(x_0; x_1, u)$. Given a part of the system

$$\begin{aligned}\dot{x}_{1,i} &= x_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(x_0, x_1, u) + cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_d} + \cdots \\ \dot{x}_{1,i+1} &= x_{0,i+2} + \sum_{k=2}^{d-1} f_{1,i+1}^{[k]}(x_0, x_1, u) + \cdots\end{aligned}\tag{8.46}$$

The following push down

$$\begin{aligned}z_{1,i+1} &= x_{1,i+1} + cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_d} \\ z_{s,t} &= x_{s,t}, \quad \text{if } (s, t) \neq (1, i+1)\end{aligned}\tag{8.47}$$

brings (8.46) to

$$\begin{aligned}\dot{z}_{1,i} &= z_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(z_0, z_1, u) + \cdots \\ \dot{z}_{1,i+1} &= z_{1,i+2} + \sum_{k=2}^{d-1} f_{1,i+1}^{[k]}(z_0, z_1, u) + \frac{d}{dt}(cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_d}) + \cdots\end{aligned}\tag{8.48}$$

Because the lowest homogeneous part of $(d/dt)(cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_d})$ is still a term of degree $[d; 0]$, it can be further pushed down. When $i = n_1$, the nonlinear term is absorbed by the feedback. Therefore, all terms of degree $[d; 0]$ can be canceled by nonlinear transformations.

Consider $f_1^{[d-1;1]}(x_0; x_1, u)$. Given a part of the system:

$$\begin{aligned}\dot{x}_{1,i-1} &= x_{1,i} + \sum_{k=2}^{d-1} \tilde{f}_{1,i-1}^{[k]}(x_0, x_1, u) + \cdots \\ \dot{x}_{1,i} &= x_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(x_0, x_1, u) + cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}}x_{1,j_d} + \cdots\end{aligned}\tag{8.49}$$

If $j_d > 1$, we can pull up the degree m term by defining

$$\begin{aligned} z_{1,i} &= x_{1,i} - cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}}x_{1,j_d-1} \\ z_{s,t} &= x_{s,t}, \quad \text{if } (s,t) \neq (1,i) \end{aligned} \quad (8.50)$$

The new system has the form:

$$\begin{aligned} \dot{z}_{1,i-1} &= z_{1,i} + \sum_{k=2}^{d-1} \tilde{f}_{1,i-1}^{[k]}(z_0, z_1, u) + cz_{1,j_1}z_{1,j_2} \cdots z_{1,j_{d-1}}z_{1,j_d-1} + \cdots \\ \dot{z}_{1,i} &= z_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(z_0, z_1, u) - \frac{d}{dt}(cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}})x_{1,j_d-1} + \cdots \end{aligned} \quad (8.51)$$

In all the new terms of degree $[d-1; 1]$, the index of the controllable factors is $j_d - 1$, which is smaller than the original index j_d . If $i = 1$, we can cancel the degree $[d-1; 1]$ term without worrying about the equation of $\dot{x}_{i-1}$. Repeat the pull up transformation until all the degree $[d-1; 1]$ terms are brought to homogeneous terms in the form $x_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}}x_{1,1}$, in which $j_d = 1$. Now, consider a part of the system

$$\begin{aligned} \dot{x}_{1,i} &= x_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{2,i}^{[k]}(x_0, x_1, u) + cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}}x_{1,1} + \cdots \\ \dot{x}_{1,i+1} &= x_{1,i+2} + \sum_{k=2}^{d-1} \tilde{f}_{2,i+1}^{[k]}(x_0, x_1, u) + \cdots \end{aligned} \quad (8.52)$$

A push down transformation

$$\begin{aligned} z_{1,i+1} &= x_{1,i+1} + cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}}x_{1,1} \\ z_{s,t} &= x_{s,t}, \quad \text{if } (s,t) \neq (1,i+1) \end{aligned} \quad (8.53)$$

yields

$$\begin{aligned} \dot{z}_{2,i} &= z_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(z_0, z_1, u) + \cdots \\ \dot{z}_{2,i+1} &= z_{1,i+2} + \sum_{k=2}^{d-1} \tilde{f}_{1,i+1}^{[k]}(z_0, z_1, u) + \frac{d}{dt}(cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}})x_{1,1} \\ &\quad + cx_{0,j_1}x_{0,j_2} \cdots x_{0,j_{d-1}}x_{1,2} + \cdots \end{aligned} \quad (8.54)$$

Repeating the push down process, all degree $[d-1; 1]$ terms are finally pushed to the equation for $\dot{x}_{1,n_1}$, where they are canceled by the feedback. Therefore, $f_1^{[d-1;1]}(x_0; x_1, u)$ can be eliminated by homogeneous transformations.

Consider $f_1^{[l_0;l_1]}(x_0; x_1, u)$ with $2 \leq l_1 \leq d$. A part of the dynamics has the form

$$\begin{aligned}\dot{x}_{1,i-1} &= x_{1,i} + \sum_{k=2}^{d-1} \tilde{f}_{1,i-1}^{[d]}(x_0, x_1, u) + \cdots \\ \dot{x}_{1,i} &= x_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[d]}(x_0, x_1, u) + c^{[l_0]}(x_0)x_{1,j_1}x_{1,j_2} \cdots x_{1,j_{l_1}} + \cdots\end{aligned}\tag{8.55}$$

The derivation of $\tilde{f}_1^{[l_0;l_1]}$ is similar to that in Section 8.2. If $j_{l_1-1} < j_{l_1} - 1$ the pull up transformation is defined by

$$\begin{aligned}z_{1,i} &= x_{1,i} - c^{[l_0]}(x_0)x_{1,j_1}x_{1,j_2} \cdots x_{1,j_{l_1-1}}x_{1,j_{l_1}-1} \\ z_{s,t} &= x_{s,t}, \quad \text{if } (s, t) \neq (1, i)\end{aligned}\tag{8.56}$$

then the dynamics becomes

$$\begin{aligned}\dot{z}_{1,i-1} &= z_{1,i} + \sum_{k=2}^{d-1} \tilde{f}_{1,i-1}^{[k]}(z_0, z_1, u) + c^{[l_0]}(z_0)z_{1,j_1}z_{1,j_2} \cdots z_{1,j_{l_1-1}}z_{1,j_{l_1}-1} + \cdots \\ \dot{z}_{1,i} &= z_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(z_0, z_1, u) \\ &\quad - c^{[l_0]}(z_0) \sum_{k=1}^{l_1-1} \frac{z_{1,j_1}z_{1,j_2} \cdots z_{1,j_{l_1-1}}z_{1,j_{l_1}-1}}{z_{1,j_k}} z_{1,j_k+1} \\ &\quad - \frac{d}{dt}(c^{[l_0]}(x_0))x_{1,j_1}x_{1,j_2} \cdots x_{1,j_{l_1-1}}x_{1,j_{l_1}-1} + \cdots\end{aligned}\tag{8.57}$$

The lowest terms in the time derivative of $c^{[l_0]}(x_0)$ are still polynomials of x_0 with the degree l_0 . As a result of the pull up, the two largest indices of z_1 in the new terms are $j_{l_1-1}, j_{l_1} - 1$ and $j_{l_1-1} + 1, j_{l_1} - 1$, which are closer together than those of the original term. If $j_{l_1-p-1} < j_{l_1-p} = j_{l_1-p+1} = \cdots j_{l_1-1} = j_{l_1} - 1$,

we define the pull up transformation by

$$\begin{aligned} z_{1,i} &= x_{0,i} - \frac{c^{[l_0]}(x_0)}{p+1} x_{1,j_1} x_{1,j_2} \cdots x_{1,j_{l_1-p-1}} x_{1,j_{l_1}}^{p+1} \\ z_{s,t} &= x_{s,t} \quad \text{for } (s,t) \neq (1,i) \end{aligned} \quad (8.58)$$

then the dynamics becomes

$$\begin{aligned} \dot{z}_{2,i-1} &= z_{1,i} + \sum_{k=2}^{d-1} \tilde{f}_{1,i-1}^{[k]}(z_0, z_1, u) \\ &\quad + \frac{c^{[l_0]}(z_0)}{p+1} z_{1,j_1} z_{1,j_2} \cdots z_{1,j_{l_1-p-1}} z_{1,j_{l_1}}^{p+1} + \cdots \\ \dot{z}_{1,i} &= z_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(z_0, z_1, u) \\ &\quad - \frac{c^{[l_0]}(z_0)}{p+1} \sum_{k=1}^{l_1-p-1} \frac{z_{1,j_1} z_{1,j_2} \cdots z_{1,j_{l_1-p-1}} z_{1,j_{l_1}}^{p+1}}{z_{1,j_k}} z_{1,j_{k+1}} \\ &\quad - \frac{d}{dt} (c^{[l_0]}(x_0)) x_{1,j_1} x_{1,j_2} \cdots x_{1,j_{l_1-p-1}} z_{1,j_{l_1}}^{p+1} + \cdots \end{aligned} \quad (8.59)$$

Notice that the two largest indices of variable $x_{1,j}$ in the new degree $[l_0; l_1]$ terms are identical. In any case, if $i = 1$ then we can still pull up and there is no $z_{1,i-1}$ dynamics to be concerned with, so a term disappears. By pulling up all the degree $[l_0; l_1]$ terms until their two largest indices of $x_{1,j}$ are identical, we obtain

$$\dot{x}_{1,i} = x_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(x_0, x_1, u) + \sum_{j=1}^{n_1+1} \epsilon_{i,j}^{[d-2]}(x_0, \tilde{x}_{1,j}) x_{1,j}^2 + \cdots \quad (8.60)$$

By pushing down we can make $\epsilon_{i,j}^{[d-2]} = 0$ for $1 \leq j \leq i+1$. Consider a piece of the dynamics,

$$\begin{aligned} \dot{x}_{1,i} &= x_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(x_0, x_1, u) + c^{[l_0]}(x_0) x_{1,j_1} x_{1,j_2} \cdots x_{1,j_{l_1}} + \cdots \\ \dot{x}_{1,i+1} &= x_{1,i+2} + \sum_{k=2}^{d-1} \tilde{f}_{1,i+1}^{[k]}(x_0, x_1, u) + \cdots \end{aligned} \quad (8.61)$$

If $1 \leq j_1 \leq j_2 \leq \cdots \leq j_{l_1} \leq n_1$, define

$$\begin{aligned} z_{1,i+1} &= x_{0,i+1} + c^{[l_0]}(x_0)x_{1,j_1}x_{1,j_2} \cdots x_{1,j_{l_1}} \\ z_{s,t} &= x_{s,t}, \quad \text{for } (s,t) \neq (1,i+1) \end{aligned} \quad (8.62)$$

yielding

$$\begin{aligned} \dot{z}_{1,i} &= z_{1,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{1,i}^{[k]}(z_0, z_1, u) + \cdots \\ \dot{z}_{1,i+1} &= z_{1,i+2} + \sum_{k=2}^{d-1} \tilde{f}_{1,i+1}^{[k]}(z_0, z_1, u) + c^{[l_0]}(z_0) \sum_{k=1}^{l_1} \frac{z_{1,j_1} z_{1,j_2} \cdots z_{1,j_{l_1}}}{z_{1,j_k}} z_{1,j_k+1} \\ &\quad \times \frac{d}{dt}(c^{[l_0]}(x_0))x_{1,j_1}x_{1,j_2} \cdots x_{1,j_{l_1}} + \cdots \end{aligned} \quad (8.63)$$

and all the other degree d terms remain unchanged. Notice that if $i+1 = n_1$ then we can absorb the degree d terms into the control using feedback. The terms in (8.60) where $1 \leq j_1 \leq j_2 \leq \cdots \leq j_{l_1} \leq i+1$ can be repeatedly pushed down and absorbed in the control. The result is the normal form of $\tilde{f}_1^{[k]}(z_0, z_1)$ in (8.45).

Consider $f_0^{[d,0]}(x_0)$. Its homological equation (8.41) is independent of the feedback. Therefore, the normal form is the same as the Poincaré normal form.

Consider $f_0^{[d-1,1]}(x_0; x_1, u)$. Given a part of the dynamics

$$\begin{aligned} \dot{x}_{0,i-1} &= \lambda_{i-1}x_{0,i-1} + \delta_{i-1}x_{0,i} + \sum_{k=2}^{d-1} \tilde{f}_{0,i-1}^{[k]}(x_0, x_1, u) + \cdots \\ \dot{x}_{0,i} &= \lambda_i x_{0,i} + \delta_i x_{0,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{0,i}^{[k]}(x_0, x_1, u) + c^{[d-1]}(x_0)x_{1,j} + \cdots \end{aligned} \quad (8.64)$$

where $2 \leq i \leq n_0$, $1 \leq j \leq n_1 + 1$. The coefficients δ_{i-1} and δ_i equal 0 or 1. If $j > 1$, then we can pull up by defining

$$\begin{aligned} z_{0,i} &= x_{0,i} - c^{[d-1]}(x_0)x_{1,j-1} \\ z_{s,t} &= x_{s,t} \quad \text{if } (s,t) \neq (0,i) \end{aligned} \quad (8.65)$$

so that

$$\begin{aligned}
 \dot{z}_{0,i-1} &= \lambda_{i-1}z_{0,i-1} + \delta_{i-1}(z_{0,i} + c^{[d-1]}(z_0)z_{1,j-1}) \\
 &\quad + \sum_{k=2}^{d-1} \tilde{f}_{1,i-1}^{[k]}(z_0, z_1, u) + \cdots \\
 \dot{z}_{0,i} &= \lambda_i z_{0,i} + \delta_i z_{0,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{0,i}^{[k]}(z_0, z_1, u) \\
 &\quad + \lambda_i c^{[d-1]}(z_0)z_{1,j-1} - \frac{d}{dt}(c^{[d-1]}(z_0))z_{1,j-1} + \cdots
 \end{aligned} \tag{8.66}$$

The new degree $[d-1; 1]$ terms have last index $1, j-1$ instead of $1, j$. We can continue to pull up until $j = 1$. The result is the normal form $\tilde{f}_1^{[d-1; 1]}$ in (8.45). If $i = 1$, the pull up cancels the $[d-1; 1]$ term if $j > 1$.

Consider $f_0^{[l_0; l_1]}(x_0; x_1, u)$ with $2 \leq l_1 \leq d$. Given a part of the system

$$\begin{aligned}
 \dot{x}_{0,i-1} &= \lambda_{i-1}x_{0,i-1} + \delta_{i-1}x_{0,i} + \sum_{k=2}^{d-1} \tilde{f}_{0,i-1}^{[k]}(x_0, x_1, u) + \cdots \\
 \dot{x}_{1,i} &= \lambda_i x_{0,i} + \delta_i x_{0,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{0,i}^{[k]}(x_0, x_1, u) \\
 &\quad + c^{[l_0]}(x_0)x_{1,j_1}x_{1,j_2} \cdots x_{1,j_{l_1}} + \cdots
 \end{aligned} \tag{8.67}$$

where $1 \leq j_1 \leq j_2 \leq \cdots \leq j_{l_1} \leq n_1 + 1$. The coefficients δ_{i-1} and δ_i equal 0 or 1. If $j_{l_1-1} < j_{l_1} - 1$, then we can pull up by defining

$$\begin{aligned}
 z_{0,i} &= x_{0,i} - c^{[l_0]}(x_0)x_{1,j_1}x_{1,j_2} \cdots x_{1,j_{l_1}-1} \\
 z_{s,t} &= x_{s,t}, \quad \text{if } (s, t) \neq (0, i)
 \end{aligned} \tag{8.68}$$

so that

$$\begin{aligned}
 \dot{z}_{0,i-1} &= \lambda_{i-1} z_{0,i-1} + \delta_{i-1}(z_{0,i} + c^{[l_0]}(z_0) z_{1,j_1} z_{1,j_2} \cdots z_{1,j_{l_1-1}}) \\
 &\quad + \sum_{k=2}^{d-1} \tilde{f}_{1,i-1}^{[k]}(z_0, z_1, u) + \cdots \\
 \dot{z}_{0,i} &= \lambda_i z_{0,i} + \delta_i z_{0,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{0,i}^{[k]}(z_0, z_1, u) \\
 &\quad + \lambda_i c^{[l_0]}(z_0) z_{1,j_1} z_{1,j_2} \cdots z_{1,j_{l_1-1}} - \frac{d}{dt}(c^{[l_0]}(z_0)) z_{1,j_1} z_{1,j_2} \cdots z_{1,j_{l_1-1}} \\
 &\quad - c^{[l_0]}(z_0) \sum_{k=1}^{l_1-1} \frac{z_{1,j_1} z_{1,j_2} \cdots z_{1,j_{l_1-1}}}{z_{1,j_k}} z_{1,j_{k+1}} + \cdots
 \end{aligned} \tag{8.69}$$

In the new $[l_0; l_1]$ terms, the two largest indices of $x_{0,j}$ are closer than before. If $j_{l_1-p-1} < j_{l_1-p} = j_{l_1-p+1} = \cdots = j_{l_1-1} = j_{l_1} - 1$ for some $p \geq 1$, define the following pull up transformation

$$\begin{aligned}
 z_{0,i} &= x_{0,i} - \frac{c^{[l_0]}(x_0)}{p+1} x_{1,j_1} \cdots x_{1,j_{l_1-p-1}} x_{1,j_{l_1-1}}^{p+1} \\
 z_{s,t} &= x_{s,t}, \quad \text{if } (s, t) \neq (0, i)
 \end{aligned} \tag{8.70}$$

Then

$$\begin{aligned}
 \dot{z}_{0,i-1} &= \lambda_{i-1} z_{0,i-1} + \delta_{i-1} \left(z_{0,i} + \frac{c^{[l_0]}(z_0)}{p+1} z_{1,j_1} \cdots z_{1,j_{l_1-p-1}} z_{1,j_{l_1-1}}^{p+1} \right) \\
 &\quad + \sum_{k=2}^{d-1} \tilde{f}_{0,i-1}^{[k]}(z_0, z_1, u) + \cdots \\
 \dot{z}_{0,i} &= \lambda_i z_{0,i} + \delta_i z_{0,i+1} + \sum_{k=2}^{d-1} \tilde{f}_{0,i}^{[k]}(z_0, z_1, u) \\
 &\quad + \lambda_i \frac{c^{[l_0]}(z_0)}{p+1} z_{1,j_1} \cdots z_{1,j_{l_1-p-1}} z_{1,j_{l_1-1}}^{p+1} \\
 &\quad - \frac{d}{dt} \left(\frac{c^{[l_0]}(x_0)}{p+1} \right) x_{1,j_1} \cdots x_{1,j_{l_1-p-1}} x_{1,j_{l_1-1}}^{p+1} \\
 &\quad - \frac{c^{[l_0]}(z_0)}{p+1} \sum_{k=1}^{l_1-p-1} \frac{z_{1,j_1} \cdots z_{1,j_{l_1-p-1}} z_{1,j_{l_1-1}}^{p+1}}{z_{1,j_k}} z_{1,j_{k+1}} + \cdots
 \end{aligned} \tag{8.71}$$

In the new $[l_0; l_1]$ terms, the last two indices of $x_{1,j}$ are equal. We repeat the pull up process until all $[l_0; l_1]$ terms have the form $c(z_0, \bar{z}_{1,j})z_{1,j}^2$. ■

8.4 Examples of Normal Form

The derivation of normal forms for specific engineering systems is not necessarily a complicated process. In this section, we introduce three examples. In each example, the normal form can be easily derived through simple transformations of push up and pull down.

8.4.1 The Normal Form of Ball and Beam

Consider the ball and beam experiment shown in Figure 8.1. The system model adopted in this section is from Hauser et al. [9]. We assume that the beam rotates around the axis at its center. The ball rolls along the beam. The control input of the system is τ , the angular acceleration of the beam. The state variables are r , the distance from the center of the ball to the axis, and θ , the angle of the beam. Let J be the moment of inertia of the beam, m be the mass of the ball, and g be the acceleration of gravity. The equations

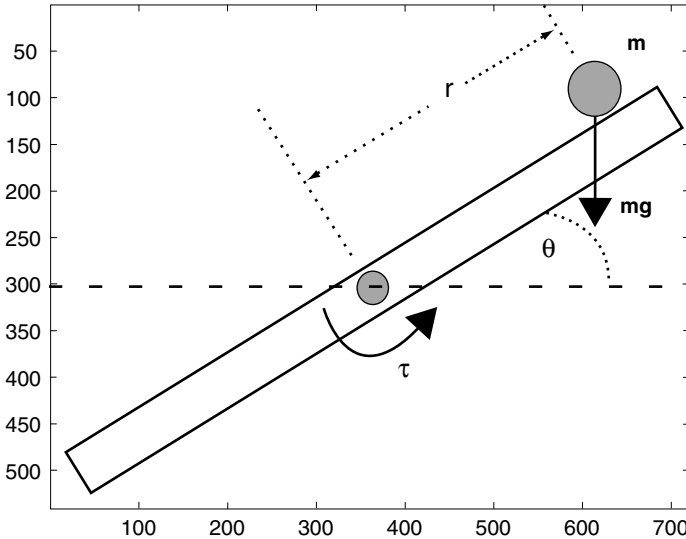


FIGURE 8.1

The configuration of ball and beam system.

of motion are

$$\begin{aligned} 0 &= \ddot{r} + g \sin \theta - r\dot{\theta}^2 \\ \tau &= (mr^2 + J)\ddot{\theta} + 2mr\dot{r}\dot{\theta} + mgr \cos \theta \end{aligned} \quad (8.72)$$

Let

$$\tau = 2mr\dot{r}\dot{\theta} + mgr \cos \theta + (mr^2 + J)u \quad (8.73)$$

This is an invertible feedback under which the system (8.72) is equivalent to

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -g \sin x_3 + x_1 x_4^2 \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= u \end{aligned} \quad (8.74)$$

where $x_1 = r$, $x_2 = \dot{r}$, $x_3 = \theta$, and $x_4 = \dot{\theta}$. The origin $(x_1, x_2, x_3, x_4) = (0, 0, 0, 0)$ is an equilibrium point of the system. The linearization of the system at the origin is

$$\begin{aligned} \delta \dot{x}_1 &= \delta x_2 \\ \delta \dot{x}_2 &= -g \delta x_3 \\ \delta \dot{x}_3 &= \delta x_4 \\ \delta \dot{x}_4 &= \delta u \end{aligned} \quad (8.75)$$

Obviously, the linearization is controllable. So, the model (8.74) of ball and beam system is linearly controllable at the origin. In the following, we derive the normal form for the system (8.74). First, we focus on the non-linear term $g \sin x_3$. We will handle the term $x_1 x_4^2$ later. Instead of dealing with the homogeneous terms separately, system (8.74) allows us to push down all the homogeneous terms in $g \sin x_3$ simultaneously. The push down transformation is

$$z_3 = -g \sin x_3 \quad (8.76)$$

after which the system becomes

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= z_3 + x_1 x_4^2 \\ \dot{z}_3 &= -g x_4 \cos x_3 \\ \dot{x}_4 &= u \end{aligned} \quad (8.77)$$

One more step of pushing down by

$$z_4 = -gx_4 \cos x_3 \quad (8.78)$$

yields

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= z_3 + x_1 x_4^2 \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= -gu \cos x_3 + gx_4^2 \sin x_3 \end{aligned} \quad (8.79)$$

If $-(\pi/2) < x_3 < (\pi/2)$ we can define an invertible feedback

$$v = -gu \cos x_3 + gx_4^2 \sin x_3$$

and then the system becomes

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= z_3 + x_1 x_4^2 \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= v \end{aligned} \quad (8.80)$$

Now, we have to deal with the term $x_1 x_4^2$ in (8.80). From (8.76) and (8.78), the inverse transformation satisfies

$$\begin{aligned} x_3 &= \arcsin\left(-\frac{z_3}{g}\right) \\ x_4 &= -\frac{z_4}{g \cos(\arcsin(-z_3/g))} \end{aligned} \quad (8.81)$$

Define $z_1 = x_1, z_2 = x_2$, (8.80) is equivalent to

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 + \frac{z_1 z_4^2}{g^2 \cos^2(\arcsin(-z_3/g))} \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= v \end{aligned} \quad (8.82)$$

However,

$$\begin{aligned}\cos^2\left(\arcsin\left(-\frac{z_3}{g}\right)\right) &= 1 - \sin^2\left(\arcsin\left(-\frac{z_3}{g}\right)\right) \\ &= 1 - \frac{z_3^2}{g^2}\end{aligned}$$

So, the system (8.82) is equivalent to

$$\begin{aligned}\dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 + \frac{z_1}{g^2 - z_3^2} z_4^2 \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= v\end{aligned}\tag{8.83}$$

This system is in normal form. Its homogeneous parts of any degree can be found in the following Taylor expansion

$$\begin{aligned}\dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 + \sum_{k=0}^{\infty} \frac{1}{g^{2k+2}} z_1 z_3^{2k} z_4^2 \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= v\end{aligned}\tag{8.84}$$

8.4.2 Engine Compressor

The second example is the Moore–Greitzer three state model of an axial flow compressor. The model is a typical example of a control system with both classical and control bifurcations. When the engine compressor is operated around the equilibrium with the maximum pressure rise, a classical bifurcation occurs in its uncontrolled dynamics. There is also a control bifurcation in the control system. On a branch of the bifurcated equilibria, the system exhibits rotating stall which can cause severe vibrations with rapid and catastrophic consequences. In the following, a model of engine compressor is introduced. Then the normal form of the model is derived at the point where rotating stall occurs.

The Moore–Greitzer model of an engine compressor described in Eveker et al. [5] is

$$\begin{aligned}\frac{dA}{d\xi} &= \frac{3\alpha H}{2W} A \left(1 - \left(\frac{\Phi}{W} - 1 \right)^2 - \frac{A^2}{4W^2} \right) \\ \frac{d\Phi}{d\xi} &= \frac{1}{l_c} \left(-\Psi + \Phi_c \left(\frac{\Phi}{W} - 1 \right) - \frac{3HA^2}{4W^2} \left(\frac{\Phi}{W} - 1 \right) \right) \\ \frac{d\Psi}{d\xi} &= \frac{1}{4l_c B^2} \left(\Phi - F_T^{-1}(\Psi) \right)\end{aligned}\quad (8.85)$$

where ξ is the scaled time. The compressor and throttle characteristics are, respectively:

$$\begin{aligned}\Phi_c(y) &= \psi_0 + H \left(1 + \frac{3}{2}y - \frac{1}{2}y^3 \right) \\ F_T^{-1}(\Psi) &= K_T \sqrt{\Psi}\end{aligned}\quad (8.86)$$

The three states in the system are A , the scaled amplitude of the rotating stall cell; Φ , the scaled annulus averaged mass flow; Ψ , the scaled annulus averaged pressure rise. The throttle parameter is K_T . When viewed as a dynamical system, K_T is a parameter and a classical bifurcation occurs at a critical value. When viewed as a control system, K_T is the control input and a control bifurcation occurs at the same critical value. The other parameters ψ_0 , H , B , α , l_c , and W are constants. More details on the meaning of the variables and the parameters are discussed in References [5, 27]. We focus on the following equilibrium point for our discussion. It is actually the stall inception point of the compressor model.

$$A_0 = 0, \quad \Phi_0 = 2W, \quad \Psi_0 = \psi_0 + 2H, \quad K_{T0} = \frac{2W}{\sqrt{\psi_0 + 2H}} \quad (8.87)$$

It is convenient to transfer the equilibrium point to the origin by the following change of coordinates

$$\Phi = \phi + 2W, \quad \Psi = \psi + \psi_0 + 2H, \quad K_T = \frac{2W}{\sqrt{\psi_0 + 2H}} + u \quad (8.88)$$

where, u is the new control input. The resulting system under the new coordinates (A, ϕ, ψ) has the following form

$$\begin{aligned}\frac{dA}{d\xi} &= \frac{3\alpha H}{2W} A \left(1 - \left(\frac{\phi}{W} + 1 \right)^2 - \left(\frac{A}{2W} \right)^2 \right) \\ \frac{d\phi}{d\xi} &= \frac{1}{l_c} \left(-\psi - \psi_0 - 2H + \Phi_c \left(\frac{\phi}{W} + 1 \right) - \frac{3HA^2}{4W^2} \left(\frac{\phi}{W} + 1 \right) \right) \\ \frac{d\psi}{d\xi} &= \frac{1}{4l_c B^2} \left(\phi + 2W - \left(\frac{2W}{\sqrt{\psi_0 + 2H}} + u \right) \sqrt{\psi + \psi_0 + 2H} \right)\end{aligned}\quad (8.89)$$

It is equivalent to

$$\begin{aligned}\frac{dA}{d\xi} &= \frac{3\alpha H}{2W} A \left(-\frac{\phi^2}{W^2} - \frac{2\phi}{W} - \frac{A^2}{4W^2} \right) \\ \frac{d\phi}{d\xi} &= \frac{1}{l_c} \left(-\psi - \frac{3H}{2W^2} \phi^2 - \frac{3H}{4W^2} A^2 - \frac{H}{2W^3} \phi^3 - \frac{3H}{4W^3} A^2 \phi \right) \\ \frac{d\psi}{d\xi} &= \frac{1}{4l_c B^2} \left(\phi + 2W - \frac{2W}{\sqrt{\psi_0 + 2H}} \sqrt{\psi + \psi_0 + 2H} + \sqrt{\psi + \psi_0 + 2H} u \right)\end{aligned}\quad (8.90)$$

The variables ψ and ϕ constitute the linearly controllable part. The normal form of the controllable part can be obtained by pushing down. Let

$$\begin{aligned}x_{0,1} &= A, \quad x_{1,1} = \phi, \\ x_{1,2} &= \frac{1}{l_c} \left(-\psi - \frac{3H}{2W^2} \phi^2 - \frac{3H}{4W^2} A^2 - \frac{H}{2W^3} \phi^3 - \frac{3H}{4W^3} A^2 \phi \right)\end{aligned}\quad (8.91)$$

The resulting system is

$$\begin{aligned}\frac{dx_{0,1}}{d\xi} &= -\frac{3\alpha H}{W^2} \left(x_{0,1} x_{1,1} + \frac{1}{8W} x_{0,1}^3 + \frac{1}{2W} x_{0,1} x_{1,1}^2 \right) \\ \frac{dx_{1,1}}{d\xi} &= x_{1,2} \\ \frac{dx_{1,2}}{d\xi} &= a(x_{0,1}, x_{1,1}, x_{1,2}) + b(x_{0,1}, x_{1,1}, x_{1,2})u\end{aligned}\quad (8.92)$$

where $a(x_{0,1}, x_{1,1}, x_{1,2}) + b(x_{0,1}, x_{1,1}, x_{1,2})u$ is defined by

$$\begin{aligned} \frac{dx_{1,2}}{d\xi} = & \frac{1}{l_c} \left(-\frac{d\psi}{d\xi} - \frac{3H}{W^2} \phi \frac{d\phi}{d\xi} - \frac{3H}{2W^2} A \frac{dA}{d\xi} \right. \\ & \left. - 3 \frac{H}{2W^3} \phi^2 \frac{d\phi}{d\xi} - \frac{3H}{4W^3} \left(2A\phi \frac{dA}{d\xi} + A^2 \frac{d\phi}{d\xi} \right) \right) \end{aligned} \quad (8.93)$$

If we define the new control input by

$$v = a(x_{0,1}, x_{1,1}, x_{1,2}) + b(x_{0,1}, x_{1,1}, x_{1,2})u \quad (8.94)$$

then we have

$$\begin{aligned} \frac{dx_{0,1}}{d\xi} &= -\frac{3\alpha H}{W^2} \left(x_{0,1}x_{1,1} + \frac{1}{8W}x_{0,1}^3 + \frac{1}{2W}x_{0,1}x_{1,1}^2 \right) \\ \frac{dx_{1,1}}{d\xi} &= x_{1,2} \\ \frac{dx_{1,2}}{d\xi} &= v \end{aligned} \quad (8.95)$$

In this system, the controllable part is in normal form. The dynamics of $x_{0,1}$ is not linearly controllable. However, this equation is already in its normal form. So, (8.95) is the normal form of the engine compressor model (8.85). Although the feedback (8.94) is complicated, only the linear and quadratic parts of a and b are critical to the bifurcations of the system [19]. Their linear and quadratic Taylor expansions are

$$\begin{aligned} a(x_{0,1}, x_{1,1}, x_{1,2}) &= -\frac{1}{4l_c^2 B^2} x_{1,1} - \frac{W}{4l_c^2 B^2 (\psi_0 + 2H)} x_{1,2} \\ &\quad - \frac{3H}{16l_c^2 B^2 W (\psi_0 + 2H)} x_{0,1}^2 - \frac{3H}{8l_c^2 B^2 W (\psi_0 + 2H)} x_{1,1}^2 \\ &\quad - \frac{3H}{l_c W^2} x_{1,1}x_{1,2} - \frac{W}{16B^2 (\psi_0 + 2H)^2} x_{1,2}^2 + O(x)^3 \\ b(x_{0,1}, x_{1,1}, x_{1,2}) &= \frac{\sqrt{\psi_0 + 2H}}{4l_c^2 B^2} - \frac{1}{8l_c B^2 \sqrt{\psi_0 + 2H}} x_{1,2} + O(x)^2 \end{aligned} \quad (8.96)$$

8.4.3 Controlled Lorenz Equation

It is known that circuit systems can be designed to approximate chaotic behavior such as the one exhibited by the Lorenz system. In References

[26, 35], the following controlled Lorenz equation is studied,

$$\begin{aligned}\dot{x} &= a(y - x) \\ \dot{y} &= cx - xz - y + u \\ \dot{z} &= xy - bz\end{aligned}\tag{8.97}$$

where a , b , and c are constant numbers. It is shown [26, 35] that several state feedbacks exist under which the closed-loop system exhibit at least three fundamentally different chaos. In the following, we use a globally invertible transformation to derive the normal form of (8.97). As a result, the entire family of control systems with the same normal form has chaotic trajectories equivalent to those found earlier [26, 35].

The transformation is simple

$$\begin{aligned}x_1 &= x \\ x_2 &= a(y - x) \\ x_0 &= z - \frac{1}{2a}x^2 \\ v &= a(cx - xz - y - ay + ax + u)\end{aligned}\tag{8.98}$$

Its inverse transformation is defined as follows

$$\begin{aligned}x &= x_1 \\ y &= x_1 + \frac{1}{a}x_2 \\ z &= x_0 + \frac{1}{2a}x_1^2 \\ u &= (1 - c)x_1 + \left(1 + \frac{1}{a}\right)x_2 + x_1x_0 + \frac{1}{2a}x_1^3 + \frac{1}{a}v\end{aligned}\tag{8.99}$$

In (8.98), x is the same as x_1 . The second equation in (8.98) is a push down. The transformation of x_0 is a pull up to cancel the term $(1/a)x_1x_2$ in the equation of $\dot{x}_0$. Under this transformation, it is easy to check

$$\begin{aligned}\dot{x}_0 &= -bx_0 + \left(1 - \frac{b}{2a}\right)x_1^2 \\ \dot{x}_1 &= x_2 \\ \dot{x}_2 &= v\end{aligned}\tag{8.100}$$

It is in normal form, with only one nonzero invariant, the coefficient of x_1^2 . If $b \neq 0$ and $2a$, the equilibrium set of the system is a parabola. The system

is linearly controllable at all its equilibrium points except for the origin. So, local control of such a system is relatively simple. However, its global behavior needs further study due to the chaotic behavior under certain state feedbacks.

8.5 Conclusions

In this chapter, normal forms of single input control systems are summarized. The system is nonlinear and the input is non-affine. The family of systems addressed in this chapter is the most general one relative to existing published normal forms of single input systems based on a similar approach. In addition, examples of normal forms are shown to illustrate the elementary transformation of push up and pull down in the derivation of normal forms. Owing to space limitation, applications of the normal forms are not addressed in the paper. However, interested readers are referred to the related publications in the references for results on bifurcation control, invariants, symmetries, and practical stabilization of nonlinear systems based on normal form approach.

References

1. V.I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, 2nd ed., Springer-Verlag, 1988.
2. J.-P. Barbot, S. Monaco, and D. Normand-Cyrot, Quadratic forms and feedback linearization in discrete time, *Int. J. Control*, 67, 567–586, 1997.
3. R.W. Brockett, Feedback invariants for nonlinear systems, in *Proceedings of the IFAC Congress*, Helsinki, 1978.
4. B. Charlet, J. Lévine, and R. Marino, On dynamic feedback linearization, *Syst. Control Lett.*, 13, 143–152, 1989.
5. K.M. Eveker, D.L. Gysling, C.N. Nett, and O.P. Sharma, Integrated control of rotating stall and surge in aeroengines, *Proc. SPIE*, 2494 (21), 21–35, 1995.
6. O.E. Fitch, The control of bifurcations with engineering applications, Ph.D. Dissertation, U.S. Naval Postgraduate School, Monterey, California, 1997.
7. B. Hamzi, J.-P. Barbot, S. Monaco, and D. Normand-Cyrot, Nonlinear discrete-time control of systems with a Naimark–Sacker bifurcation, *Syst. Control Lett.*, 44, 245–258, 2001.
8. B. Hamzi, W. Kang, and J.-P. Barbot, Analysis and control of Hopf bifurcations, *SIAM J. Control Optim.*, 42 (6), 2200–2220, 2004.

9. J. Hauser, S. Sastry, and P. Kokotović, Nonlinear control via approximate input–output linearization: the ball and beam example, in *Proceedings of the IEEE Conference on Decision and Control*, Tampa, Florida, December, 1989.
10. B. Jakubczyk and W. Respondek, On linearization of control systems, *Bull. Acad. Polon. Sci. Ser. Math.*, 28, 517–522, 1980.
11. L.R. Hunt and R. Su, Linear equivalents of nonlinear time varying systems, in *Proceedings of the MTNS*, Santa Monica, CA, 119–123, 1981.
12. W. Kang, Extended controller normal form, invariants and dynamic feedback linearization of nonlinear control systems, Dissertation, University of California, Davis, California, 1991.
13. W. Kang and A.J. Krener, Extended quadratic controller normal form and dynamic feedback linearization of nonlinear systems, *SIAM J. Control Optim.*, 30, 1319–1337, 1992.
14. W. Kang, Approximate linearization of nonlinear control systems, *Syst. Control Lett.*, 23, 43–52, 1994.
15. W. Kang, Quadratic normal forms of nonlinear control systems with uncontrollable linearization, in *Proceedings of the 34th IEEE CDC*, 608–612, 1995.
16. W. Kang, Extended controller form and invariants of nonlinear control systems with a single input, *J. Math. Syst., Estim. Control*, 6, 27–51, 1996.
17. W. Kang, Bifurcation and normal form of nonlinear control systems: Part I, *SIAM J. Control Optim.*, 36, 193–212, 1998.
18. W. Kang, Bifurcation and normal form of nonlinear control systems: Part II, *SIAM J. Control Optim.*, 36, 213–232, 1998.
19. W. Kang, Bifurcation control via state feedback for systems with a single uncontrollable mode, *SIAM J. Control Optim.*, 38, 1428–1452, 2000.
20. W. Kang, M. Xiao, and I. Tall, Controllability and local accessibility – a normal form approach, *IEEE Trans. Autom. Control*, 48 (10), 1724–1736, 2003.
21. A.J. Krener, Approximate linearization by state feedback and coordinate change, *Syst. Control Lett.*, 5, 181–185, 1984.
22. A.J. Krener, Normal forms for linear and nonlinear systems, in *Differential Geometry, the Interface Between Pure and Applied Mathematics*, M. Luksik, C. Martin, and W. Shadwick (eds.), Contemporary Mathematics, Vol. 68, American Mathematical Society, Providence, pp. 157–189, 1986.
23. A.J. Krener, S. Karahan, M. Hubbard, and R. Frezza, Higher order linear approximations to nonlinear control systems, in *Proceedings of the IEEE Conference On Decision and Control*, Los Angeles, pp. 519–523, 1987.
24. A.J. Krener and L. Li, Normal forms and bifurcations of discrete-time nonlinear control systems, *SIAM J. Control Optim.*, 40, 1697–1723, 2002.
25. A.J. Krener, W. Kang, and D. Chang, Control bifurcations, *IEEE Trans. Automat. Control*, 49 (8), 1231–1246, 2004.
26. J. Lü and G. Chen, A new chaotic attractor coined, *Int. J. Bifurcation Chaos*, 12, 659–661, 2002.
27. F.E. McCaughan, Bifurcation analysis of axial flow compressor stability, *SIAM J. Appl. Math.*, 20, 1232–1253, 1990.
28. W. Respondek and I.A. Tall, How many symmetries does admit a nonlinear single-input control system around an equilibrium?, in *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, Florida, pp. 1795–1800, 2001.

29. W. Respondek and I.A. Tall, Nonlinearizable single-input control systems do not admit stationary symmetries, *Syst. Control Lett.*, 46, 1–16, 2002.
30. W. Respondek, Symmetries and minimal flat output of nonlinear control systems, in *New Trends in Nonlinear Dynamics and Control, and Their Applications*, W. Kang, M. Xiao, and C. Borges (eds.), Springer-Verlag, Berlin, 2003.
31. I. Tall and W. Respondek, Normal forms and invariants of nonlinear single-input systems with noncontrollable linearization, in *Proceedings of the IFAC Symposium on Nonlinear Control Systems*, St. Petersburg, Russia, July, 2001.
32. I.A. Tall and W. Respondek, Normal forms of two-inputs nonlinear control systems, in *Proceedings of the 41th CDC*, Las Vegas, USA, 2002.
33. I. A. Tall, Normal forms of multi-inputs nonlinear systems with controllable linearization, in *Lecture Notes in Control and Information Sciences*, W. Kang, M. Xiao, and C. Borges (eds.), Springer-Verlag, Berlin, 2003.
34. I.A. Tall and W. Respondek, Feedback classification of nonlinear single-input control systems with controllable linearization: normal forms, canonical forms, and invariants, *SIAM J. Control Optim.*, 41, 1498–1531, 2003.
35. T. Ueta and G. Chen, Bifurcation analysis of Chen's equation, *Int. J. Bifurcation Chaos*, 10, 1917–1931, 2000.

9

Observability Bifurcations: Application to Cryptography

J.-P. Barbot, I. Belmouhoub, and L. Boutat-Baddas

CONTENTS

9.1	Introduction	377
9.2	Some Recalls on Observability	381
9.3	Observability Normal Form	384
9.3.1	Continuous Time Case	384
9.3.2	Discrete Time Case	388
9.4	Unknown Input Observer	392
9.5	Synchronization of Chaotic Systems	393
9.5.1	Continuous Time Transmission: The Chua Circuit	393
9.5.2	Discrete Time Transmission: Burger's Map	397
9.5.2.1	Ciphering	398
9.5.2.2	Deciphering	398
9.6	Rössler Map for Hyperchaotic Synchronization	399
9.6.1	Step-by-Step Delayed Reconstructor Design	401
9.7	Conclusions	403
	Acknowledgements	403
9.8	Appendix	404
9.8.1	Proof of Proposition 1	404
9.8.2	Proof of Theorem 2	404
	References	407

9.1 Introduction

This chapter deals with the use of the normal forms to analyze observability bifurcations. This concept of normal form was introduced in the

context of control theory by Kang and Krener [28]. Here, we present and use the normal form in order to analyze observability bifurcations (more precisely observability singularities manifolds) and then apply these to the design of chaotic ciphering and deciphering. Before dealing with this particular application, some basic notions on the system synchronization will be recalled. In a recent paper, Tauleigne et al. [56] have highlighted the difference between synchronization and syntonization (also called unidirectional synchronization in the literature [13]). More precisely, the syntonization is a special class of synchronization, where one system is the master and the other is the slave. The slave system must totally or partially reproduce the behavior of the master. The syntonization problem¹ was very well defined as a problem of observer design by Nijmeijer and Mareels [43]. However, the case of bidirectional synchronization is more difficult to characterize using the usual concepts of control theory.

To this purpose, consider the following two identical bidirectionally coupled systems shown in Figure 9.1.

Choosing $x_1 = V_{c1}$, $x_2 = i_1$, $x_3 = V_{c2}$, $x_4 = i_2$ as system state and taking $C_1 = C_2 = C$ and $L_1 = L_2 = L$, the following state equations are obtained:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} \frac{-1}{CR} & \frac{-1}{C} & \frac{1}{CR} & 0 \\ \frac{1}{L} & 0 & 0 & 0 \\ \frac{1}{CR} & 0 & \frac{-1}{CR} & \frac{-1}{C} \\ 0 & 0 & \frac{1}{L} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad (9.1)$$

The state matrix in (9.1) has two purely imaginary eigenvalues and also two eigenvalues with negative real parts. Therefore, in this simple case, it is possible to predict the behavior of both bidirectionally coupled systems, only with eigenvalue and eigenvector arguments. In the directions of the purely imaginary eigenvalues, the system behavior is that of two synchronous oscillators. In the directions of the eigenvalues with negative real part the amplitude asymptotically goes to zero. Nevertheless, in order to emphasize the physical meaning, consider the following Lyapunov function:

$$V = \frac{1}{2}L(x_2^2 + x_4^2) + \frac{1}{2}C(x_1^2 + x_3^2) \quad (9.2)$$

¹In this chapter, the word synchronization refers to either a bidirectional synchronization or a syntonization (which can be viewed as a unidirectional synchronization).

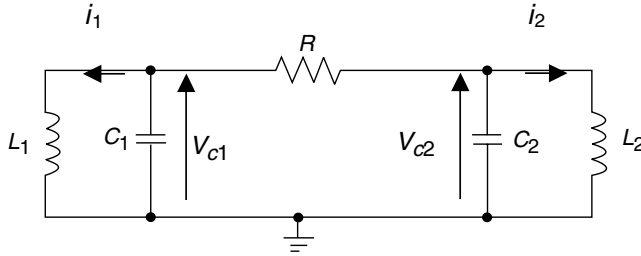


FIGURE 9.1
Bidirectional synchronization.

The derivative of V along the trajectories of (9.1) is

$$\dot{V} = \frac{-1}{R}(x_1 - x_3)^2 \quad (9.3)$$

So the system (9.1) is proved to be Lyapunov stable [31]. To invoke the LaSalle theorem [31, 40], the set S of equi-Lyapunov is defined as:

$$S = \{x \in R^4 / \dot{V} = 0\} \quad (9.4)$$

For each bounded initial condition, all behaviors of (9.1) are guaranteed to be bounded according to (9.3). Therefore, all behaviors of (9.1) asymptotically converge to the invariant subset² of S . Here, the invariant subset of S is:

$$IS = \{x \in R^4 / x_1 = x_3 \quad \text{and} \quad x_2 = x_4\} \quad (9.5)$$

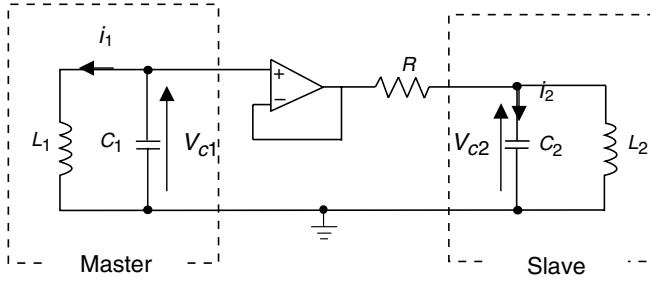
The set IS corresponds to all the perfectly synchronized behaviors of (9.1).³ Now, consider the unidirectional synchronization scheme shown in Figure 9.2.

Setting again the same state vector $x = (V_{c1}, i_1, V_{c2}, i_2)$ and taking $C_1 = C_2 = C$ and $L_1 = L_2 = L$, the following state equations are obtained:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} 0 & \frac{-1}{C} & 0 & 0 \\ \frac{1}{L} & 0 & 0 & 0 \\ \frac{1}{CR} & 0 & \frac{-1}{CR} & \frac{-1}{C} \\ 0 & 0 & \frac{1}{L} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad (9.6)$$

²Subset invariant with respect to the system of bidirectionally coupled oscillators.

³Condition $x_2 = x_4$ is obtained with $\dot{V} = 0$.

**FIGURE 9.2**

Unidirectional synchronization.

The state matrix has two purely imaginary eigenvalues and two eigenvalues with negative real parts. This second system (9.6) is also easy to analyze thanks to eigenvalues and eigenvectors arguments. In this case, however, some observability arguments can be used. The system (9.6) may be divided into two parts: the master dynamics (x_1 and x_2) and the slave (the observer) dynamics (x_3 and x_4). From linear control theory arguments, it can be proved that the master system with x_1 as output is observable. Moreover, defining $e_1 = x_1 - x_3$ and $e_2 = x_2 - x_4$ as observation error on x_1 and x_2 , respectively, we obtain:

$$\begin{pmatrix} \dot{e}_1 \\ \dot{e}_2 \end{pmatrix} = \begin{pmatrix} \frac{-1}{CR} & \frac{-1}{C} \\ \frac{1}{L} & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad (9.7)$$

The system (9.7) has two eigenvalues with negative real parts. This very simple example highlights the interest of the observer approach in the case of unidirectional synchronization.⁴ The previous two examples seem to show that the analysis of the unidirectional synchronization is easier than that of the bidirectional one.

Moreover, in the unidirectional synchronization case, only stability and observability bifurcation can appear and for bidirectional synchronization stability, observability and *controllability* bifurcation can appear.

This chapter only deals with unidirectional synchronization. To improve the safety of data transmission, an unidirectional synchronization with at least one observability bifurcation is designed. Moreover, to deeply hide the message a new ciphering scheme is proposed. This new scheme will be referred to as an “inclusion method” [1, 5, 6, 10] in contrast to the classical one, the “addition method.”⁵

⁴All generalizations starting from a particular case must be realized cautiously.

⁵This method is also called chaotic masking in the literature [13].

The loss of the linear observability property at one point or on a submanifold is called observability bifurcation. In many previous chapters, the notion of bifurcation was first dedicated to stability properties, and the concept of normal form was introduced to analyze the stability bifurcations [47]. The main idea behind this concept is to highlight the influences of the dominant terms with respect to a considered local property. Moreover, each normal form characterizes one and only one equivalent class. So, the structural properties of the normal form are also the same as those of each system in the corresponding equivalence class. Other works have also recalled the usefulness of normal form to analyze the controllability bifurcations [3, 18–20, 27, 28, 35, 36, 55].

In some recent papers [6, 10], we have introduced a new class of homogeneous transformations, by diffeomorphism and output injection, which is used to obtain an equivalence relation. Then normal form for each equivalence class can be chosen. All these works will be recalled subsequently in this chapter, and the fact that all systems of the same equivalence class have the same observability properties will be highlighted. Specifically, some observer design will be proposed on the basis of the normal form structure. Moreover, the concept of observability matching condition [45] will be recalled on the basis of the normal form structure to deal with the inclusion method. The last part of this chapter is dedicated to the synchronization of hyperchaotic systems [6]: as in some recent papers [5, 26, 42], we highlight the fact that the Pyragas conjecture⁶ [49] may be overcome and that the synchronization of hyperchaotic system is possible with only one output. Moreover, some observability bifurcations are added to improve the data transmission security in our discrete-time example.

9.2 Some Recalls on Observability

One of the first definitions and characterizations of nonlinear observability was given in the well-known paper of Hermann and Krener [22], where the concept of local weak observability was introduced and the observability rank condition was given. In their paper [22], observability and controllability were studied with the same tools as of differential geometry [41]. As with the linear case, some direct links between observability and controllability may be found. After this pioneering paper many works on nonlinear observability followed [8, 53]. An important fact, pointed out

⁶The minimal number of outputs to synchronize two hyperchaotic systems is equal to the number of positive Lyapunov exponents of the master system.

in the 1980s, was the loss of observability due to an inappropriate input. Consequently, the characterization of appropriate input (universal input) with respect to nonlinear observability [16] was an important challenge. Since then, much research has been done on the design of nonlinear observers. From our point of view, one of the first significant theoretical and practical contributions to the subject was the linearization by output injection proposed by Krener and Isidori [37] for a single output system and by Krener and Respondek [38] for the multi-output case (see also Xia and Gao [60]). From these works and other ones dealing with structural analysis [7, 17, 24, 30, 46, 52], important work on nonlinear observer design followed. Different techniques were studied: high gain [17, 32], backstepping [29, 50], extended Luenberger [9], Lyapunov approach [57], sliding mode [4, 15, 45, 54, 61], numerical differentiator [14], etc.

Some observer designs partially or totally use the notion of detectability. This concept will be used and highlighted in this paper in the context of observability bifurcation (see also the paper of Krener and Xiao [39]).

In this section, we only recall concepts of weak observability, observability rank condition, and linearization by output injection.

Consider the following system:

$$\dot{x} = f(x), \quad y = h(x) \quad (9.8)$$

where vector fields $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ are assumed to be smooth with $f(0) = 0$ and $h(0) = 0$. The observability problem arises as follows: can we estimate the current state $x(t)$ from past observations $y(s)$, $s \leq t$. An algorithm that solves this problem is called an observer. Before designing an observer for the system (9.8), we must check whether it verifies some conditions as it is weakly locally observable.

DEFINITION 1 The observability codistribution of the system (9.8) denoted dO , is the codistribution defined for $x \in \Xi \subset \mathbb{R}^n$ (with Ξ a neighborhood of an equilibrium point) by

$$dO(x) = \begin{pmatrix} dh_j \\ dL_f h_j \\ dL_f^2 h_j \\ \vdots \\ dL_f^i h_j \\ \vdots \end{pmatrix} \quad \text{with } j \in \{1, \dots, m\} \text{ and } i \in N^* \quad (9.9)$$

The main theorem concerning weak local observability [22] is the following:

THEOREM 1

Let dO be the observability codistribution associated to the system (9.8). If

$$\dim(dO(x)) = n \quad (9.10)$$

(9.8) is locally observable at x .

REMARK 1

- If (9.8) is a linear system, then we recover the classical rank condition.
- If (9.8) is a non-autonomous system, then the observability codistribution depends on the input and the weak local observability also depends on the input (u) and their derivatives. The input which preserves the observability properties is called universal input [16].

Motivated by the consideration that it is always possible to cancel all independent parts constituted only by the input and the output in the estimated error, the observer linearization problem was born. Is it possible to find in a neighborhood U of 0 in $\mathbb{R}^n$ a change of state coordinates $z = \theta(x)$ such that dynamic (9.8) is linear driven by nonlinear output injection:

$$\dot{z} = Az - \beta(y) \quad (9.11)$$

where $\beta: \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a smooth vector field. Note that the output injection term $\beta(y)$ is cancelled in the observation error dynamic for system (9.11). The diffeomorphism θ must satisfy the first-order partial differential equation:

$$\frac{\partial \theta}{\partial x}(x)f(x) = A\theta(x) - \beta(h(x)) \quad (9.12)$$

Krener and Isidori showed [37] that equation (9.12) has a solution if and only if the following two conditions are satisfied:

1. The codistribution span $\{dh, dL_f h, \dots, dL_f^{n-1} h\}$ is of rank n at 0
2. $[\tau, ad_f^k \tau] = 0$ for all $k = 1, 3, \dots, 2n - 1$, where τ is the unique solution vector fields of $[(dh)^T, (dL_f h)^T, \dots, (dL_f^{n-1} h)^T]^T \tau = [0, 0, \dots, 1]^T$

REMARK 2

- Condition 1 implies that linear tangent approximation is observable.

- Generally, condition 2 is difficult to verify for noninvolutivity reason. As the noninvolutivity of a distribution implies its nonintegrability (see the Frobenius theorem [25, 44, 58]) Krener [34] introduced the concept of approximated integrability (see also [12, 21] for the same notion around a manifold) which is one of the key points of controllability and observability normal forms.

9.3 Observability Normal Form

In this section, we first present the normal form for a system linearly observable; then we present the normal form for a system with one linear unobservable mode and these in both continuous and discrete-time case. For the sake of compactness, the proof of the main results is given in Appendix A.

9.3.1 Continuous Time Case

Consider a nonlinear single input single output (SISO) system:

$$\dot{\xi} = f(\xi) + g(\xi)u, \quad y = C\xi \quad (9.13)$$

where, vector fields $f, g: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ are assumed to be real analytic, such that $f(0) = 0$ and $y \in \mathbb{R}$.

Setting $A = (\partial f / \partial \xi)(0)$ and $B = g(0)$ around the equilibrium point $\xi_e = 0$, the system can be rewritten in the following form:

$$\dot{z} = Az + Bu + f^{[2]}(z) + g^{[1]}(z)u + O^{[3]}(z, u), \quad y = Cz \quad (9.14)$$

where $f^{[2]}(z) = [f_1^{[2]}(z), \dots, f_n^{[2]}(z)]^T$ and $g^{[1]}(z) = [g_1^{[1]}(z), \dots, g_n^{[1]}(z)]^T$ for all $1 \leq i \leq n$, $f_i^{[2]}(z)$ and $g_i^{[1]}(z)$ are, respectively, homogeneous polynomials of degree 2 and 1 in z .

DEFINITION 2

1. The component $f^{[2]}(z) + g^{[1]}(z)u$ is the quadratic part of system (9.14).
2. Consider a second system:

$$\dot{x} = Ax + Bu + \bar{f}^{[2]}(x) + \bar{g}^{[1]}(x)u + O^{[3]}(x, u), \quad y = Cx \quad (9.15)$$

We say that system (9.14) whose quadratic part is $f^{[2]}(z) + g^{[1]}(z)u$, is QEMOI (Quadratically Equivalent Modulo an Output Injection) to

system (9.15) whose quadratic part is $\tilde{f}^{[2]}(x) + \tilde{g}^{[1]}(x)u$, if there exists an output injection:

$$\beta^{[2]}(y) + \gamma^{[1]}(y)u \quad (9.16)$$

and a diffeomorphism of the form:

$$x = z - \Phi^{[2]}(z) \quad (9.17)$$

which carries $f^{[2]}(z) + g^{[1]}(z)u$ to $\tilde{f}^{[2]}(x) + \tilde{g}^{[1]}(x)u + [\beta^{[2]}(y) + \gamma^{[1]}(y)u]$. Here $\Phi^{[2]}(z) = [\Phi_1^{[2]}(z), \dots, \Phi_n^{[2]}(z)]^T$, $\beta^{[2]}(y) = [\beta_1^{[2]}(y), \dots, \beta_n^{[2]}(y)]^T$ and for all $1 \leq i \leq n$, $\Phi_i^{[2]}(z)$ and $\beta_i^{[2]}(y)$ are homogeneous polynomials of degree 2 in, respectively, z and y and $\gamma^{[1]}(y) = [\gamma_1^{[1]}(y), \dots, \gamma_n^{[1]}(y)]^T$ where $\gamma_i^{[1]}(y)$ is a homogeneous polynomial of degree 1 in y .

3. If $\tilde{f}^{[2]}(x) = 0$ and $\tilde{g}^{[1]}(x) = 0$, we say that system (9.14) is quadratically linearizable modulo an output injection.

REMARK 3

- If $((\partial f/\partial x)(0), C)$ is observable, then one can transform system (9.13) to the following form:

$$\begin{cases} \dot{z} = A_0 z + B_0 u + f^{[2]}(z) + g^{[1]}(z)u + O^{[3]}(z, u) \\ y = z_1 = C_0 z \end{cases} \quad (9.18)$$

$$\text{with: } A_0 = \begin{pmatrix} a_1 & 1 & 0 & \cdots & 0 \\ a_2 & 0 & 1 & 0 & \cdots \\ \vdots & 0 & \vdots & \vdots & 0 \\ a_{n-1} & 0 & \cdots & 0 & 1 \\ a_n & 0 & \cdots & \cdots & 0 \end{pmatrix}, \quad B_0 = \begin{pmatrix} b_1 \\ \vdots \\ \vdots \\ \vdots \\ b_n \end{pmatrix}$$

- If $((\partial f/\partial x)(0), C)$ has one unobservable real mode, then one can transform system (9.13) to the following form:

$$\begin{cases} \dot{\tilde{z}} = A_{obs} \tilde{z} + B_{obs} u + \tilde{f}^{[2]}(z) + \tilde{g}^{[1]}(z)u + O^{[3]}(z, u) \\ \dot{z}_n = \alpha_n z_n + \sum_{i=1}^{n-1} \alpha_i z_i + b_n u + f_n^{[2]}(z) + g_n^{[1]}(z)u + O^{[3]}(z, u) \\ y = z_1 = C_{obs} \tilde{z} \end{cases} \quad (9.19)$$

$$\text{with: } \tilde{z} = \begin{pmatrix} z_1 \\ \vdots \\ \vdots \\ \vdots \\ z_{n-1} \end{pmatrix}, \quad z = [\tilde{z}^T, z_n]^T,$$

$$A_{obs} = \begin{pmatrix} a_1 & 1 & 0 & \cdots & 0 \\ a_2 & 0 & 1 & 0 & \cdots \\ \vdots & 0 & \vdots & \vdots & 0 \\ a_{n-2} & 0 & \cdots & 0 & 1 \\ a_{n-1} & 0 & \cdots & \cdots & 0 \end{pmatrix}, \quad B_{obs} = \begin{pmatrix} b_1 \\ \vdots \\ \vdots \\ \vdots \\ b_{n-1} \end{pmatrix}$$

Throughout the chapter, the output is always taken to be equal to the first state component. Consequently, the diffeomorphism ($x = z - \Phi^{[2]}(z)$) is such that $\Phi_1^{[2]}(z) = 0$.

PROPOSITION 1 [10]

System (9.14) is QEMOI to system (9.15), if and only if the following two homological equations are satisfied:

$$\begin{aligned} 1. \quad & A\Phi^{[2]}(z) - \frac{\partial \Phi^{[2]}}{\partial z}Az = \bar{f}^{[2]}(z) - f^{[2]}(z) + \beta^{[2]}(z_1) \\ 2. \quad & -\frac{\partial \Phi^{[2]}}{\partial z}B = \bar{g}^{[1]}(z) - g^{[1]}(z) + \gamma^{[1]}(z_1) \end{aligned} \tag{9.20}$$

where $(\partial \Phi^{[2]}/\partial z)Az := (\partial \Phi_1^{[2]}(z)/\partial z)Az, \dots, (\partial \Phi_n^{[2]}(z)/\partial z)Az^T$ and $\partial \Phi_i^{[2]}(z)/\partial z$ is the Jacobian matrix of $\Phi_i^{[2]}(z)$ for all $1 \leq i \leq n$.

The proof of this proposition is presented in the appendix.

Now we can give the normal form associated to QEMOI relation. First, we recall the theorem introduced in Reference [10] for a system which is linearly observable.

THEOREM 2

There is a quadratic diffeomorphism and an output injection which transform system (9.18) in the following observability normal form:

$$\left\{ \begin{array}{l} \dot{x}_1 = a_1 x_1 + x_2 + b_1 u + \sum_{i=2}^n k_{1i} x_i u + O^{[3]}(x, u) \\ \vdots \\ \dot{x}_{n-1} = a_{n-1} x_1 + x_n + b_{n-1} u + \sum_{i=2}^n k_{(n-2)i} x_i u + O^{[3]}(x, u) \\ \dot{x}_n = a_n x_1 + b_n u + \sum_{j=2}^n h_{ij} x_i x_j + h_{1n} x_1 x_n \\ \quad + \sum_{i=2}^n k_{(n-1)i} x_i u + O^{[3]}(x, u) \end{array} \right. \quad (9.21)$$

The proof is reported in the appendix.

COROLLARY 1

There is a quadratic diffeomorphism and an output injection which transform system (9.19) in the following observability normal form:

$$\left\{ \begin{array}{l} \dot{x}_1 = a_1 x_1 + x_2 + b_1 u + \sum_{i=2}^n k_{1i} x_i u + O^{[3]}(x, u) \\ \vdots \\ \dot{x}_{n-2} = a_{n-2} x_1 + x_{n-1} + b_{n-2} u + \sum_{i=2}^n k_{(n-2)i} x_i u + O^{[3]}(x, u) \\ \dot{x}_{n-1} = a_{n-1} x_1 + b_{n-1} u + \sum_{j=2}^n h_{ij} x_i x_j + h_{1n} x_1 x_n \\ \quad + \sum_{i=2}^n k_{(n-1)i} x_i u + O^{[3]}(x, u) \\ \dot{x}_n = \alpha_n x_n + \sum_{i=1}^{n-1} \alpha_i x_i + b_n u + \alpha_n \Phi_n^{[2]}(x) + \sum_{i=1}^{n-1} \alpha_i \Phi_i^{[2]}(x) \\ \quad - \frac{\partial \Phi_n^{[2]}}{\partial \tilde{x}} A_{obs} \tilde{x} + f_n^{[2]}(x) + \sum_{i=2}^n k_{ni} x_i u + O^{[3]}(x, u) \end{array} \right. \quad (9.22)$$

The proof is similar to the previous one, for more details, see Reference [10].

REMARK 4

1. If, for some index $i \in [1, n]$ we have $h_{in} x_i \neq 0$, then we can recover, at least locally, all state components.
2. If we have some $k_{in} \neq 0$ then with an appropriate choice of input u (universal input [16]) we can have quadratic observability.
3. Thus, the local quadratic observability is principally given by the dynamic $\dot{x}_{n-1}$. In the case where conditions 1 and 2 are not verified, then we can use coefficient α_n to study the detectability propriety. Then, we have three cases:
 - (a) If $\alpha_n < 0$ then the state x_n is detectable
 - (b) If $\alpha_n > 0$ then x_n is unstable, and consequently undetectable

- (c) If $\alpha_n = 0$ we can use the center manifold theory to analyze stability or instability of x_n and consequently its detectability or undetectability.

Remembering the well-known Poncaré–Dulac theorem⁷ it is expected that resonant terms may appear, then we have:

If $\Phi_n^{[2]}(x)$ verifies the following equation:

$$\alpha_n \Phi_n^{[2]}(x) + \sum_{i=1}^{n-1} \alpha_i \Phi_i^{[2]}(x) = \frac{\partial \Phi_n^{[2]}}{\partial x} A_{obs} x - f_n^{[2]}(x) + \beta_n^{[2]}(x_1) \quad (9.23)$$

then quadratic terms in $\dot{x}_n$ are cancelled, which is not the case, in general, for arbitrary α_n and a_i . Nevertheless, this condition is less restrictive than the usual one, thanks to the output injection $\beta_n^{[2]}(x_1)$.

9.3.2 Discrete Time Case

Now, let us consider a discrete time nonlinear SISO system:

$$\xi^+ = f(\xi, u), \quad y = C\xi \quad (9.24)$$

where, ξ is the state of the system and ξ (resp. ξ^+) denote $\xi(k)$ (resp. $\xi(k+1)$). The vector field $f: U \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ and the function $h: M \subset \mathbb{R}^n \rightarrow \mathbb{R}$ are assumed to be real analytic, such that $f(0,0) = 0$.

As for the continuous time case, we recall the observability normal form for a linear observable system and for a system with one linear unobservable mode. We apply, as usual, a second-order Taylor expansion around the equilibrium point.

Thus the system is rewritten as:

$$\begin{cases} z^+ = Az + Bu + F^{[2]}(z) + g^{[1]}(z)u + \gamma^{[0]}u^2 + O^3(z, u) \\ y = Cz \end{cases} \quad (9.25)$$

with $A = (\partial f / \partial x)(0,0)$, $B = (\partial f / \partial u)(0,0)$ and where: $F^{[2]}(z) = [F_1^{[2]}(z), \dots, F_n^{[2]}(z)]^T$ and $g^{[1]}(z) = [g_1^{[1]}(z), \dots, g_n^{[1]}(z)]^T$. Now, we define the QEMOI for a discrete time system given by (9.25).

⁷The condition for resonance is: there exists an eigenvalue L_i such that $L_i = \sum_{j=1}^n m_j L_j$ with $m_j \in \mathbb{N}$ and $\sum_{j=1}^n m_j \geq 2$ [59].

DEFINITION 3 The system:

$$z^+ = Az + Bu + F^{[2]}(z) + g^{[1]}(z)u + \gamma^{[0]}u^2 + O^3(z, u), \quad y = Cz \quad (9.26)$$

is said to be quadratically equivalent to the system:

$$\begin{cases} x^+ = Ax + Bu + \bar{F}^{[2]}(x) + \bar{g}^{[1]}(x)u + \bar{\gamma}^{[0]}u^2 \\ \quad + \beta^{[2]}(y) + \alpha^{[1]}(y)u + \tau^{[0]}u^2 + O^3(x, u) \\ y = Cx \end{cases} \quad (9.27)$$

modulo the output injection:

$$\beta^{[2]}(y) + \alpha^{[1]}(y)u + \tau^{[0]}u^2 \quad (9.28)$$

if there exists a diffeomorphism of the form:

$$x = z - \Phi^{[2]}(z) \quad (9.29)$$

which transforms the quadratic part of (9.26) into the one of (9.27).

REMARK 5

The output injection (9.28) is different from the one defined in (9.16) for the continuous-time case. This is due to the fact that the vector field composition does not preserve the linearity in “ u ”; so we have to consider the term $\tau^{[0]}u^2$ in (9.28).

In the next proposition, we give the necessary and sufficient conditions for QEMOI:

PROPOSITION 1

System (9.26) is QEMOI to system (9.27), if and only if there exist $(\Phi^{[2]}, \beta^{[2]}, \alpha^{[1]}, \gamma^{[0]})$ which satisfy the following homological equations:

- (i) $F^{[2]}(x) - \bar{F}^{[2]}(x) = \Phi^{[2]}(Ax) - A\Phi^{[2]}(x) + \beta^{[2]}(x_1)$
- (ii) $g^{[1]}(x) - \bar{g}^{[1]}(x) = \widehat{\Phi}^{[2]}(Ax, B) + \alpha^{[1]}(x_1)$
- (iii) $\gamma^{[0]} - \bar{\gamma}^{[0]} = \Phi^{[2]}(B) + \tau^{[0]}$

Proof of Proposition 1 is the same as that of Proposition 1 in continuous time case but with some additional technical difficulties due to the composition operator; see Reference [6] for complete proof.

Now, to apply our study to a system linearly observable and a system with one unobservable mode, let us consider:

- The system (9.13) where the pair (A, C) is observable. Then there is a linear change of coordinates ($z = T\xi$) and a Taylor expansion which transforms the system (9.13) into the following form:

$$\begin{cases} z^+ = A_0 z + B_0 u + F^{[2]}(z) + g^{[1]}(z)u + \gamma^{[0]}u^2 + O^3 \\ y = z_1 = C_0 z \end{cases} \quad (9.30)$$

where A_0 , B_0 , and C_0 are as defined in (9.18).

- The system (9.13) where the pair (A, C) has one unobservable mode. Then there is a linear change of coordinates ($z = T\xi$) and a Taylor expansion which transforms the system (9.13) into the following form:

$$\begin{cases} \tilde{z}^+ = A_{obs} \tilde{z} + B_{obs} u + \tilde{F}^{[2]}(z) + \tilde{g}^{[1]}(z)u + \tilde{\gamma}^{[0]}u^2 + O^3 \\ z_n^+ = \eta z_n + \sum_{i=1}^{n-1} \lambda_i z_i + b_n u + F_n^{[2]}(z) + g_n^{[1]}(z)u + \gamma_n^{[0]}u^2 + O^3 \\ y = C_{obs} z \end{cases} \quad (9.31)$$

where A_{obs} , B_{obs} , and C_{obs} are as defined in (9.19).

The quadratic normal form associated with system (9.30) is given in the following theorem.

THEOREM 3

The discrete time observability normal form with respect to the quadratic equivalence modulo an output injection of the system (9.30) is:

$$\begin{cases} x_1^+ = a_1 x_1 + x_2 + b_1 u + \sum_{i=2}^n k_{1i} x_i u \\ \vdots \\ x_{n-1}^+ = a_{n-1} x_1 + x_n + b_{n-1} u + \sum_{i=2}^n k_{(n-2)i} x_i u \\ x_n^+ = a_n x_1 + b_n u + \sum_{j>i=1}^n h_{ij} x_i x_j + h_{nn} x_n^2 + \left(\sum_{i=2}^n k_{ni} x_i \right) u \end{cases} \quad (9.32)$$

Proof of the theorem can be found in Reference [6].

The following corollary is given for system with one linearly unobservable mode.

COROLLARY 2

The discrete time observability normal form with respect to the quadratic equivalence modulo an output injection of the system (9.31) is:

$$\begin{cases} x_1^+ = a_1 x_1 + x_2 + b_1 u + \sum_{i=2}^n k_{1i} x_i u \\ \vdots \\ x_{n-2}^+ = a_{n-2} x_1 + x_{n-1} + b_{n-2} u + \sum_{i=2}^n k_{(n-2)i} x_i u \\ x_{n-1}^+ = a_{n-1} x_1 + b_{n-1} u + \sum_{j>i=1}^n h_{ij} x_i x_j + h_{nn} x_n^2 + \left(\sum_{i=2}^n k_{i(n-1)} x_i \right) u \end{cases} \quad (9.33)$$

Moreover, for the linear unobservable mode the general dynamic is:

$$x_n^+ = \eta x_n + \sum_{i=1}^{n-1} \lambda_i x_i + b_n u + \sum_{(i,j) \in I, j \neq 1} l_{ij} x_i x_j + \left(\sum_{i=2}^n k_{ni} x_i \right) u \quad (9.34)$$

(This last dynamic may be simplified in some particular cases, see Reference [6].)

For the proof, see Reference [6].

REMARK 6

- The normal form (9.33) is structurally different from the controllability discrete time normal form, given in References [18, 19], in the last state dynamic x_n^+ . For the observability analysis, the main structural information is not in the x_n^+ dynamic but in the previous state evolution (x_i^+ for $n-1 \geq i \geq 1$). The terms $\lambda_i x_i$, $b_n u$, $F_n^{[2]}(x)$, $g_n^{[1]}(x)u$ are only important in the case of detectability analysis when $\eta = \pm 1$.
- Thanks to the quadratic term $k_{n(n-1)} x_n u$ in the normal form described earlier, it is possible to restore observability with a well chosen input u .
- In the normal form, let us consider more closely the observability singularity's (here we consider system without input) by isolating the terms in x_n which appear in the $(n-1)$ th line, as follows:

$$\sum_{j>i=1}^{n-1} h_{ij} x_i x_j + \left(\sum_{i=1}^n h_{in} x_i \right) x_n \quad (9.35)$$

we can deduce the manifold of local unobservability :

$$S_n = \left\{ \sum_{i=1}^{n-1} h_{in} x_i + 2h_{nn} x_n = 0 \right\}.$$

9.4 Unknown Input Observer

The observer design for a system with unknown input has been studied [23, 61] and numerous relevant applications of such approaches have been given. In this chapter, we propose to find a new application domain for the unknown input observer design. More precisely, we propose a new type of secure data transmission based on chaotic synchronization. For this we have to recall and give some particular concepts of an observer for a system with unknown input.

Roughly speaking, in a linear context, the problem of observer design for a system with unknown input is solved as follows:

Assume an observable system with two outputs and one unknown input such that at least one derivative of the output is a function of the unknown input (i.e. C_1G or C_2G different from zero),

$$\dot{x} = Ax + Bu + G\omega, \quad y_1 = C_1x, \quad y_2 = C_2x$$

Then to design an observer, we choose a new output as $y_{\text{new}} = \phi(y_1, y_2)$ and find observation error dynamics which are orthogonal to the unknown input vector. Unfortunately, this kind of design cannot be applied to a system with only one output (the case considered in this chapter). Nevertheless, it is possible with a step by step procedure to design an observer for such a system. Obviously, there are some restrictive conditions on the system allowing to solve this problem [45, 61]. Now, let us consider the nonlinear analytic system:

$$\dot{x} = f(x) + g(x)u, \quad y = h(x) \quad (9.36)$$

where vector fields f and $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ are assumed to be smooth with $f(0) = 0$ and $h(0) = 0$. Now, we can give a particular constraint to solve this problem. The unknown input observer design is solvable locally around $x = 0$ for system (9.36) if:

- $\text{span}\{dh, dL_f h, \dots, dL_f^{n-1} h\}$ is of rank n at $x = 0$
- $((dh)^T (dL_f h)^T \dots (dL_f^{n-1} h)^T)^T g = (0 \dots 0 \star)^T$ (observability matching condition, OMC)

$\star$ means a non-null term almost everywhere in the neighborhood of $x = 0$.

Sketch of proof: Setting $z_1 = h$, $z_2 = L_f h, \dots, z_n = L_f^{n-1} h$, we have

$$\dot{z}_1 = z_2, \quad \dot{z}_2 = z_3, \dots, \dot{z}_{n-1} = z_n, \quad \dot{z}_n = \tilde{f}(z) + \tilde{g}(z)u \quad (9.37)$$

Then under classical boundary assumptions, it is possible to design a step by step sliding mode observer for the system (9.37) such that we recover, in finite time, all state components and the unknown input.

DEFINITION 4 The discrete time observability matching condition (DOMC) is:

$$((dh)^T(df \circ h)^T \cdots (df_o^{n-1} \circ h)^T)^T g = (0 \cdots 0 \star)^T$$

where $\circ$ denotes the usual composition function and f_o^j denotes the function f composed j times.

9.5 Synchronization of Chaotic Systems

We now propose a new encoding algorithm based on a chaotic system synchronization but for which we also have an observability bifurcation. In both continuous and discrete time cases, the message is included in the system structure and the observability matching condition is required. Then the general transmission scheme for the inclusion method⁸ is as shown in Figure 9.3.

The inclusion method must be compared with the classical one, the so-called addition method (Figure 9.4), which is also a slave–master scheme but where the message is only masked by the chaotic signal and included in the structure of the chaotic system.

9.5.1 Continuous Time Transmission: The Chua Circuit

To illustrate our purpose, consider the well-known Chua circuit with a variable inductor (see Figure 9.5). The circuit contains linear resistors (R, R_0), a single nonlinear resistor ($f(v_1)$), and three linear energy-storage elements: a variable inductor (L) and two capacitors (C_1, C_2). The state equations for

⁸The problem of recovering the message in the inclusion method may be interpreted as a left invertible problem.

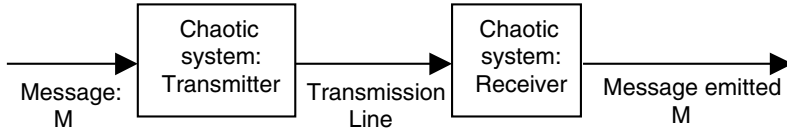


FIGURE 9.3
Inclusion method.

the circuit are as follows:

$$\begin{aligned}
 \dot{x}_1 &= \frac{-1}{C_1 R} (x_1 - x_2) + \frac{f(x_1)}{C_1} \\
 \dot{x}_2 &= \frac{1}{C_2 R} (x_1 - x_2) + \frac{x_3}{C_2} \\
 \dot{x}_3 &= -x_4 (x_2 + R_0 x_3) \\
 \dot{x}_4 &= \sigma
 \end{aligned} \tag{9.38}$$

with: $y = x_1 = v_1$, $x_2 = v_2$, $x_3 = i_3$, $x_4 = 1/L(t)$, $x = (x_1, x_2, x_3, x_4)^T$, and $f(x_1) = G_b x_1 + 0.5(G_a - G_b)(|x_1 + E| - |x_1 - E|)$.

Moreover x_1 is the output and $x_4 = 1/L$ is the only state component directly influenced by σ , an unknown bounded function. The variation of L is the information to pass on the receiver. Moreover, we assume that there exist K_1 and K_2 such that $|x_4| < K_1$ and $|dx_4/dt| < K_2$, this means that the information signal and its variation are bounded.

This system has one unobservable real mode, and using the linear change of coordinates $z_1 = x_1$, $z_2 = (x_1/C_2 R) + (x_2/C_1 R)$, $z_3 = (x_3/C_1 C_2 R)$, and

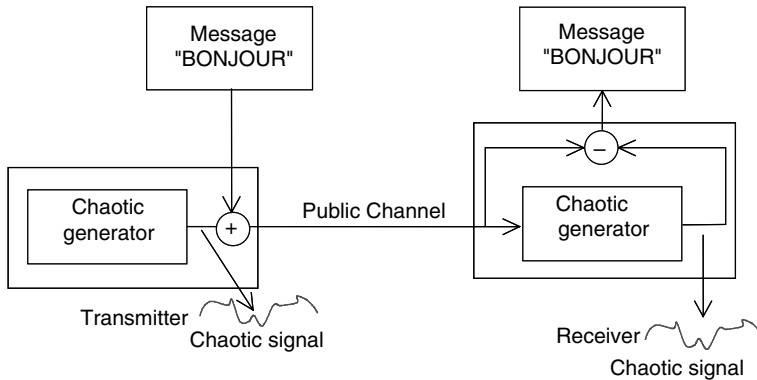


FIGURE 9.4
Addition method.

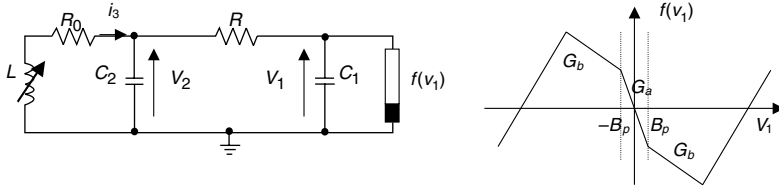


FIGURE 9.5
The Chua circuit.

$z_4 = x_4$ we obtain:

$$\begin{aligned}
 \dot{z}_1 &= \frac{-(C_1 + C_2)}{C_1 C_2 R} z_1 + z_2 + \frac{f(x_1)}{C_1} \\
 \dot{z}_2 &= z_3 + \frac{f(x_1)}{C_1 C_2 R} \\
 \dot{z}_3 &= \frac{z_1 z_4}{C_2^2 R} - \frac{z_2 z_4}{C_2} - R_0 z_3 z_4 \\
 \dot{z}_4 &= \sigma
 \end{aligned} \tag{9.39}$$

Equations (9.39) are in observability normal form [10] with $\alpha = 0$ and resonant terms are $h_{22} = h_{23} = 0$, $h_{14} = 1/C_2^2 R$, $h_{24} = 1/C_2$, and $h_{34} = -R_0$.

Moreover, the system verifies the OMC [2, 61] with respect to σ and as nonsmooth output injection $(f(x_1)/C_1, f(x_1)/(C_1 C_2 R), 0, 0)^T$. From the normal form (9.39) we conclude that the observability singularity manifold is $M_0 = \{z(z_1/(C_2^2 R)) - (z_2/C_2) - R_0 z_3 = 0\}$. Taking into account this singularity, it is possible to design the following step by step sliding mode observer (given here in the original coordinate):

$$\begin{aligned}
 \frac{d\hat{x}_1}{dt} &= \frac{1}{C_1} \left(\frac{\hat{x}_2 - y}{R} - f(y) \right) + \lambda_1 \text{sign}(y - \hat{x}_1) \\
 \frac{d\hat{x}_2}{dt} &= \frac{1}{C_2} \left(\frac{y - \tilde{x}_2}{R} + \hat{x}_3 \right) + E_1 \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2) \\
 \frac{d\hat{x}_3}{dt} &= \hat{x}_4 (-\tilde{x}_2 - R_0 \tilde{x}_3) + E_2 \lambda_3 \text{sign}(\tilde{x}_3 - \hat{x}_3) \\
 \frac{d\hat{x}_4}{dt} &= E_3 \lambda_4 \text{sign}(\tilde{x}_4 - \hat{x}_4)
 \end{aligned} \tag{9.40}$$

with the following conditions: if $\hat{x}_1 = x_1$ then $E_1 = 1$ else $E_1 = 0$, similarly if $[\hat{x}_2 = \tilde{x}_2$ and $E_1 = 1]$ then $E_2 = 1$ else $E_2 = 0$ and finally if $[\hat{x}_3 = \tilde{x}_3$ and $E_2 = 1]$ then $E_3 = 1$ else $E_3 = 0$. Moreover, to take into account the observability

singularity manifold M_0 , ($\tilde{x}_2 + R_0\tilde{x}_3 = 0$), we set $E_s = 1$ if $\tilde{x}_2 + R_0\tilde{x}_3 \neq 0$ else $E_s = 0$. And by definition we take:

$$\begin{aligned}\tilde{x}_2 &= \hat{x}_2 + E_1 C_1 R \lambda_1 \text{sign}(y - \hat{x}_1) \\ \tilde{x}_3 &= \hat{x}_3 + E_2 C_2 \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2) \\ \tilde{x}_4 &= \hat{x}_4 - \frac{E_3 E_s}{(\tilde{x}_2 + R_0\tilde{x}_3 - 1 + E_s)} \lambda_3 \text{sign}(\tilde{x}_3 - \hat{x}_3)\end{aligned}\tag{9.41}$$

Then the observation error dynamics ($e = x - \hat{x}$) are:

$$\begin{aligned}\frac{de_1}{dt} &= \frac{e_2}{C_1 R} - \lambda_1 \text{sign}(x_1 - \hat{x}_1) \\ \frac{de_2}{dt} &= \frac{e_3}{C_2} - \lambda_2 \text{sign}(x_2 - \hat{x}_2) \\ \frac{de_3}{dt} &= -(x_2 + R_0 x_3) e_4 - \lambda_3 \text{sign}(x_3 - \hat{x}_3) \\ \frac{de_4}{dt} &= \sigma - E_s \lambda_4 \text{sign}(\tilde{x}_4 - \hat{x}_4)\end{aligned}\tag{9.42}$$

The proof of observation error convergence is in Reference [11].

REMARK 7

In practice, we add some low pass filter on the auxiliary components $\tilde{x}_i$ and we set $E_i = 1$ for $i \in \{1, 2, 3\}$, not exactly when we are on the sliding manifold but when we are close enough. Similarly, $E_s = 0$ when we are close to the singularity, not only when we are on it.

To illustrate the efficiency of the method, we chose to transmit the following message: $0.1 \sin(100t)$. The message was introduced in the Chua circuit as follows: $L(t) = L + 0.1L \sin(100t)$ with: $L = 18.8$ mH.

In Figure 9.6, if we set $E_s = 0$ on a big neighborhood of the singularity manifold ($x_2 + R_0 x_3$), for a long time we lose the information on x_4 . We notice that the convergence of the state $\hat{x}_4$ of the observer, towards x_4 of the system of origin (9.38), depends on the choice of E_s (see first two curves of Figure 9.6). To have good convergence it is necessary to take $E_s = 0$ on a very small neighborhood of the singularity manifold ($x_2 + R_0 x_3$), as we can notice in the last two curves of Figure 9.6.

In any case, these simulations confirm that the resonant terms ($-x_4 x_2 - R_0 x_4 x_3$) $\neq 0$ allow us to obtain the message.

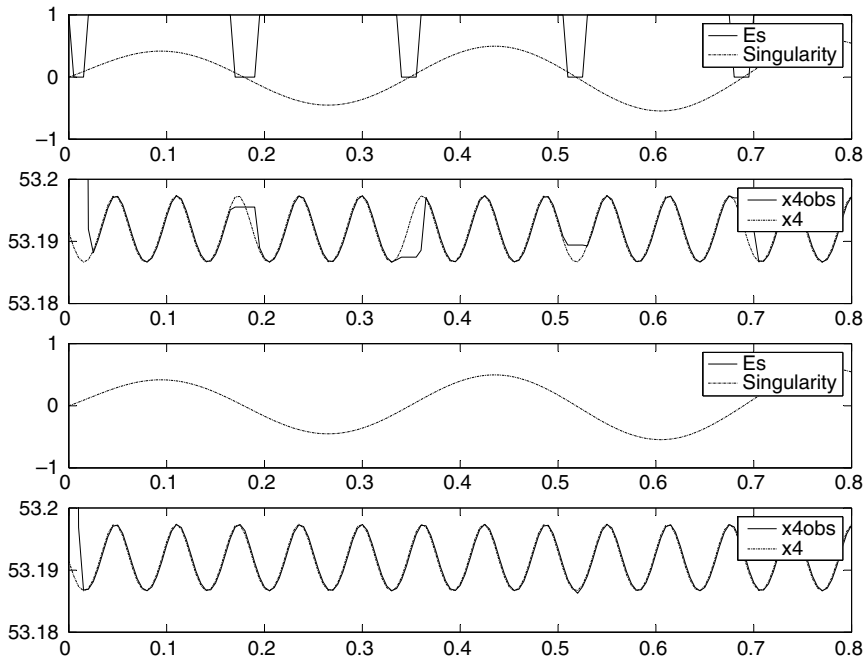


FIGURE 9.6
Observation errors.

9.5.2 Discrete Time Transmission: Burger's Map

Information is getting more and more digitized, treated, and exchanged by computers nowadays; thus we think it is of primary importance to study systems in discrete time. Here, we study the following discrete time chaotic system called the Burger map [33]:

$$\begin{cases} x_1^+ = (1 + a)x_1 + x_1x_2 \\ x_2^+ = (1 - b)x_2 - x_1^2 \end{cases} \quad (9.43)$$

where a and b are two real parameters. We assume that we can measure the state x_1 , so we have $y = x_1$ as the output of the system. This system is the normal form of:

$$\begin{cases} z_1^+ = (1 + a)z_1 + z_1z_2 \\ z_2^+ = (1 - b)(z_2 + bz_1z_2) - z_1^2 \end{cases} \quad (9.44)$$

obtained, modulo $(-y^2)$, by applying to it the change of coordinates $x = z - \Phi^{[2]}(z)$; where the diffeomorphism $\Phi^{[2]}$ is: $\Phi_1^{[2]}(z) = 0$ and $\Phi_2^{[2]}(z) = z_1 z_2$.

9.5.2.1 Cipherring

Now consider the Burger's map, and let m be the confidential message. Moreover, only the output $y = x_1$ is transmitted to the receiver via a public channel. Then, the transmitter will have the form:

$$\begin{cases} x_1^+ = (1 + a)x_1 + x_1 x_2 \\ x_2^+ = (1 - b)x_2 - x_1^2 + m \end{cases} \quad (9.45)$$

The key of this secure communication consists in the knowledge of the parameters a and b . The fact that the message should be the last information to reach in the system constitutes a necessary and sufficient condition to recover the message by the construction of a suitable observer. It is the so-called DOMC.

9.5.2.2 Deciphering

To decipher the message we construct the observer:

$$\begin{cases} \hat{x}_1^+ = (1 + b)y + y\hat{x}_2 \\ \hat{x}_2^+ = (1 - a)\hat{x}_2 - y^2 \end{cases} \quad (9.46)$$

The observer design consists in recovering the linearly unobservable states (i.e., x_2) in the observer, with the knowledge of y :

x_2 recovering: For the sake of causality, we extract x_2 from e_1 , at the iteration $(k - 1)$, which we approximate by $\tilde{x}_2^-$; so $\tilde{x}_2^- = e_1/y^-$, for $y \neq 0$. Consequently, when $y = 0$ this leads to a singularity. However, we overcome this problem by forcing $\tilde{x}_2$ to take its last remembered value when $y = 0$.

Correction of $\hat{x}_2^-$: By correction, we mean to replace $\hat{x}_2$ by $\tilde{x}_2$ in the prediction equation of $\hat{x}_2$, then we have: $\hat{x}_{2C}^- = (1 - b)\tilde{x}_2^{--} - (y^-)^2$.

The message m recovering: We have $x_2^- = \tilde{x}_2^- = (1 - b)\tilde{x}_2^{--} - (y^-)^2 + m^{--}$. It is now possible to extract m with two delays from $\tilde{e}_2$ as: $\tilde{e}_2^- = \tilde{x}_2^- - \hat{x}_{2C}^- = m^{--}$. Which means that $\tilde{e}_2(k - 1) = m(k - 2)$. So we have to wait two steps (these correspond to the necessary steps of the synchronization).

9.6 Rössler Map for Hyperchaotic Synchronization

To further improve the transmission security, we use hyperchaotic systems which features offer more guaranties for communications security. For that reason the inclusion method was adapted for *hyperchaotic systems*, from which results the “discrete time hyperchaotic-cryptography by inclusion method” (DTHCIM).

In this section, we present the DTHCIM for a three-dimensional discrete time hyperchaotic system (see the behavior of the Rossler map in Figure 9.7) introduced by Rössler [51]:

$$\begin{cases} x_1^+ = a_1 x_1(1 - x_1) + a_2 x_2 \\ x_2^+ = b_1 [(1 - b_2 x_1)(x_2 + b_3) - 1] (1 - b_4 x_3) \\ x_3^+ = c_1 x_3(1 - x_3) - c_2(1 - b_2 x_1)(x_2 + b_3) \end{cases} \quad (9.47)$$

System (9.47) can be represented in the generic form:

$$x^+ := f(x, p) \quad (9.48)$$

where $x := (x_1, x_2, x_3)^T \in \mathbb{R}^3$ represents the state vector evaluated at step k (i.e., $x(k)$), so $x^+ = x(k+1)$. The vector $p \in \mathbb{R}^8$ denotes the parameter vectors of system (9.47), $p = (a_1, a_2, b_1, b_2, b_3, b_4, c_1, c_2)^T$. Let us define $f(x, p) := (f_1(x_1, x_2, p), f_2(x_1, x_2, x_3, p), f_3(x_1, x_2, x_3, p))^T$ such that $f: \mathbb{R}^3 \times \mathbb{R}^8 \rightarrow \mathbb{R}^3$ is the analytic hyperchaotic generator of (9.47).

The Rössler map possesses six stationary points with four complex and two real points and has at least two positive Lyapunov exponents for

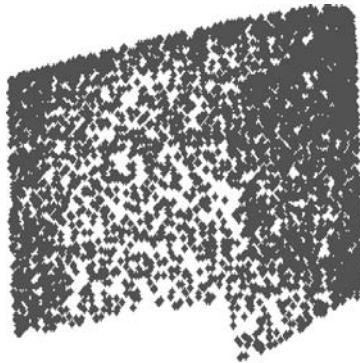


FIGURE 9.7
Rossler map phase portrait.

each real point. Starting from the hyperchaotic generator (9.48), a secured transmission scheme is constructed. The secret key is made up of part or totality of the hyperchaotic system parameters. In this case, we consider the parameter vector $(a_2, b_1, b_4)^T$ as a key.

The transmitter is represented by:

$$\begin{cases} x^+ = f(x, p) + g(x, p)u \\ y = h(x) \end{cases} \quad (9.49)$$

where x, p and vector-field f are as defined in (9.48). Vector field $h: \mathbb{R}^3 \rightarrow \mathbb{R}$ is such that $h(x) = x_1$, and vector field $g: \mathbb{R}^3 \times \mathbb{R}^8 \rightarrow \mathbb{R}^3$ will be specified next. We recall that input $u \in \mathbb{R}$ is the confidential information to transmit. Hence, system (9.49) is considered to be a SISO system with u as unknown input and x_1 as the output.

So, following DOMC we hide the confidential message u in the third dynamic x_3^+ of the transmitter (i.e., $g(x, p) = (0, 0, 1)^T$). This ensures recovering the message almost everywhere. It will be proved next that for such injection the DOMC holds.

The first point of the DOMC to be verified is whether the observation matrix J is of rank 3 almost everywhere. So, we have:

$$\det(J) = -a_2^2 b_1 b_4 [(1 - b_2 x_1)(x_2 + b_3) - 1]$$

It ensues that $\det(J) \neq 0$ for all $x \in \mathbb{R}^3/S$ such that a_2, b_1 , and b_4 are non-null.

Where singularity manifold S is given by:

$$S = \{x \in \mathbb{R}^3 \text{ such that } osg := (1 - b_2 x_1)(x_2 + b_3) - 1 = 0\}$$

So, the matrix J is of rank 3 everywhere except for S .

We conclude that system (9.49) is, at least almost everywhere, non-linearly observable. However, it possesses, for $x \in S$, one observability singularity in x_3 direction.

As for the second condition, a matrix-vector product is computed:

$$\begin{aligned} J \cdot g &= ((dh)^T, (d(f \circ h))^T, (d(f_o^2 \circ h))^T)^T g \\ &= (0, 0, -a_2 b_1 b_4 osg)^T \end{aligned}$$

We deduce that $J \cdot g = (0, 0, \star)^T$, where $\star$ is a non-null function of $x \in \mathbb{R}^3/S$.

For $x \in S$, $J \cdot g = (0, 0, 0)^T$. Consequently, through the bifurcation manifold S , there is a temporary loss of information. The message u cannot be

recovered each time the system trajectory interacts with the bifurcation manifold.

This problem is bypassed by stopping the emission of the confidential information each time an observability bifurcation is detected and by emitting any data instead. An original manner for data encryption was proposed, but the main task still is the decryption. For this purpose, a “step-by-step delayed reconstructor” is developed to recover all information.

9.6.1 Step-by-Step Delayed Reconstructor Design

To realize the decryption, the receiver is designed as a *decision and control bloc* (see Figure 9.3), able to reconstruct the original data: the “step-by-step delayed reconstructor”.

The message u is reconstructed in three steps and with three delays, from the delayed output and the delayed reconstructed dynamics of the transmitter.

REMARK 8

For convenience, we use the following notations:

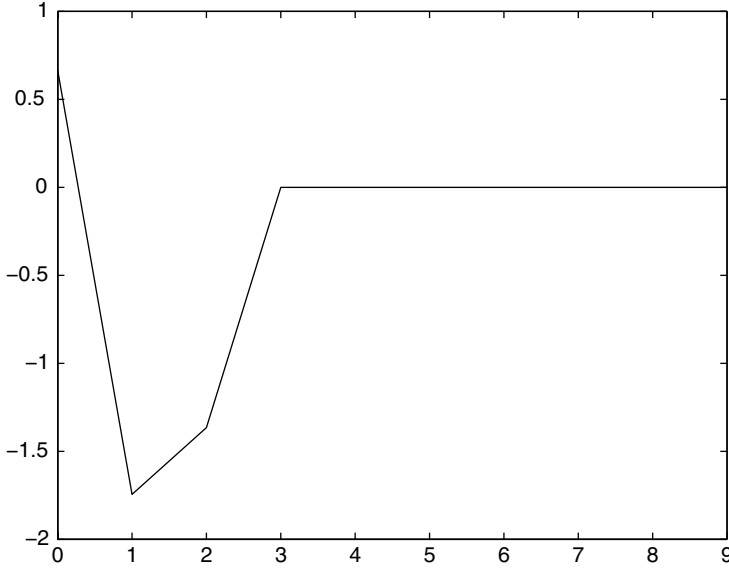
1. $\tilde{x}_i$ denotes the transmitter reconstructed dynamic: x_i for $1 \leq i \leq n$ (here $n = 3$).
2. $\forall k \in \mathbb{N}$, $(x_i)^{j-} = x_i(k - j)$, $(\tilde{x}_l)^{j-} = \tilde{x}_l(k - j)$, $y^{j-} = y(k - j)$ and w^{j-} denotes $u(k - j)$ for $1 \leq i, j \leq n$ and $2 \leq l \leq n$.

The *step-by-step delayed reconstructor*, consists in constructing step by step, transmitter dynamics with some delays, such that each reconstructed dynamic at the k th iteration arises in the construction of the next dynamic at the $(k - 1)$ th iteration until the last one, which contains information at the $(k - 3)$ th iteration. This explains why the information u^{3-} is extracted instead of u . So the reconstructor structure can be assimilated as a “go-back machine”.

In each iteration k , we extract $u(k - 3)$ as follows:

First step: Consists computing the second dynamic of the delayed reconstructor with the knowledge of the transmitter output y uniquely. It is obtained by a simple inversion: $\tilde{x}_2^- = (y - a_1 y^- (1 - y^-) / a_2$.

Second step: This step consists in finding the third delayed reconstructor dynamics. In this step are the constructed delayed dynamics: $\tilde{x}_2^-$ and $(\tilde{x}_2)^{2-}$ and the delayed output y^{2-} . This reconstruction points out the observability bifurcation highlighted in the previous observability study: $(osg)^{2-} = (1 - b_2 y^{2-})(\tilde{x}_2)^{2-} + b_3 - 1$.

**FIGURE 9.8**

Three steps convergence of signal observation error.

This singularity is bypassed by forcing $(\tilde{x}_3)^{2-}$ to take its last buffered value when $\tilde{x}^{2-} \in S$:

$$(\tilde{x}_3)^{2-} = \begin{cases} \frac{b_1(osg)^{2-} - \tilde{x}_2^-}{b_1 b_4 (osg)^{2-}} & \text{for } \tilde{x}^{2-} \in \mathfrak{N}^3/S \\ (\tilde{x}_3)^{3-} & \text{for } \tilde{x}^{2-} \in S \end{cases}$$

Last step: After the reconstruction of the last dynamic, the message (a square signal of amplitude 10^{-2}) is extracted with three delays (see Figure 9.8) by a simple differentiation under the knowledge of delayed dynamics $(\tilde{x}_3)^{2-}$, $(\tilde{x}_3)^{3-}$, $(\tilde{x}_2)^{3-}$ constructed in the previous steps and the delayed output y^{3-} :

$$\begin{aligned} u^{3-} = & (\tilde{x}_3)^{2-} - c_1(\tilde{x}_3)^{3-}(1 - (\tilde{x}_3)^{3-}) \\ & + c_2(1 - b_2 y^{3-})((\tilde{x}_2)^{3-} + b_3) \end{aligned}$$

For simulations we have:

$$\begin{aligned} p = & (a_1, a_2, b_1, b_2, b_3, b_4, c_1, c_2)^T \\ = & (3.78, 0.2, 0.1, 2, 0.35, 1.9, 3.8, 0.05)^T \end{aligned}$$

REMARK 9

The secured transmission techniques (based on chaotic synchronization) until now [48] are based on an “output injection scheme,” under the assumption of linear-observability of the hyperchaotic generator. This induces the necessity to transmit as many nonlinearities as Lyapounov exponents (at least two nonlinearities in the hyperchaotic case). The cryptographic scheme developed earlier, answers the aforementioned issue. In fact, it is based on an observable hyperchaotic map which is linearly unobservable. Moreover, on the transmission line, only one information materialized by the output of the transmitter is found.

9.7 Conclusions

In this chapter, we have first dealt with the observability bifurcation and the normal form associated with the observability properties. After that, we have highlighted the usefulness of the observability normal form for designing a secure data transmission by unidirectional synchronization of two chaotic systems. Moreover, using the concept of observers with unknown input and the so-called observer matching condition the inclusion method was recalled. This method is a ciphering method where the message is included in the structure of the chaotic transmitter. Finally, in the discrete time case, for systems that may have an observability bifurcation of dimension greater than one, we have proposed a new discrete time observer design. The proposed design requires some specific conditions: observability and observability matching condition. However, in contrast with the Pyragas conjecture, only the linear output of the transmitter is required to ensure the synchronization and a correct deciphering. So, this chapter underlines the efficiency of the observer design approach in the unidirectional synchronization. Nevertheless, as it was pointed out in section 9.1, the case of bidirectional synchronization is more difficult to analyze. We think that this case must be analyzed thanks to some energy concepts (passivity), some invariant concepts (LaSalle), and analysis of all types of bifurcations (stability, controllability, and observability) must be taken into account in this case.

Acknowledgements

The authors would like to thank M. Djemai, R. Tauleigne, and D. Boutat for their help and fruitful discussions.

9.8 Appendix

9.8.1 Proof of Proposition 1

Consider system (9.14) and let $x = z - \Phi^{[2]}(z)$ then:

$$\begin{aligned}\dot{x} = & Az + Bu + f^{[2]}(z) + g^{[1]}(z)u + O^3(z, u) \\ & - \frac{\partial \Phi^{[2]}}{\partial z}(Az + Bu + f^{[2]}(z) + g^{[1]}(z)u + O^3(z, u))\end{aligned}$$

and comparing its quadratic part with the following system:

$$\begin{aligned}\dot{x} = & A(z - \Phi^{[2]}(z)) + Bu + \bar{f}^{[2]}(z - \Phi^{[2]}(z)) + \bar{g}^{[1]}(z - \Phi^{[2]}(z))u \\ & + \beta^{[2]}(z_1) + \gamma^{[1]}(z_1)u + O^3(z - \Phi^{[2]}(z), u)\end{aligned}$$

we obtain:

$$\begin{aligned}Az - A\Phi^{[2]}(z) + \bar{f}^{[2]}(z) + \bar{g}^{[1]}(z)u + \beta^{[2]}(z_1) + \gamma^{[1]}(z_1)u \\ Az + f^{[2]}(z) + g^{[1]}(z)u - \frac{\partial \Phi^{[2]}}{\partial z}(Az + Bu)\end{aligned}$$

which gives the homological equations stated in Proposition 1:

1. $A\Phi^{[2]}(z) - \frac{\partial \Phi^{[2]}}{\partial z}Az = \bar{f}^{[2]}(z) - f^{[2]}(z) \quad \text{modulo } \beta^{[2]}(y)$
2. $-\frac{\partial \Phi^{[2]}}{\partial z}B = \bar{g}^{[1]}(z) - g^{[1]}(z) \quad \text{modulo } \gamma^{[1]}(y)$

9.8.2 Proof of Theorem 2

Setting $\bar{f}^{[2]}(z) = 0$, $\bar{g}^{[1]}(z) = 0$, $\beta^{[2]}(z_1)$, and $\gamma^{[1]}(z_1)$ in homological equations of Proposition 1 we have:

1. $A_{obs}\Phi^{[2]}(z) - \frac{\partial \Phi^{[2]}}{\partial z}A_{obs}z = -f^{[2]}(z) + \beta^{[2]}(z_1)$
2. $-\frac{\partial \Phi^{[2]}}{\partial z}B = -g^{[1]}(z) + \gamma^{[1]}(z_1)$

Thanks to the condition $\Phi_1^{[2]}(z) = 0$ (i.e., $y = z_1 = x_1$) and to the structure of A_{obs} , the first homological equation gives:

$$\begin{aligned}\Phi_2^{[2]}(z) &= -f_1^{[2]}(z) + \beta_1^{[2]}(z_1) \\ \Phi_3^{[2]}(z) &= \sum_{i=1}^{n-1} \frac{\partial \Phi_2^{[2]}}{\partial z_i} (a_i z_1 + z_{i+1}) + \frac{\partial \Phi_2^{[2]}}{\partial z_n} a_n z_1 - f_2^{[2]}(z) + \beta_2^{[2]}(z_1) \\ \Phi_4^{[2]}(z) &= \sum_{i=1}^{n-1} \frac{\partial \Phi_3^{[2]}}{\partial z_i} (a_i z_1 + z_{i+1}) + \frac{\partial \Phi_3^{[2]}}{\partial z_n} a_n z_1 - f_3^{[2]}(z) + \beta_3^{[2]}(z_1) \\ &\vdots \quad \quad \quad \vdots \\ \Phi_n^{[2]}(z) &= \sum_{i=1}^{n-1} \frac{\partial \Phi_{n-1}^{[2]}}{\partial z_i} (a_i z_1 + z_{i+1}) + \frac{\partial \Phi_{n-1}^{[2]}}{\partial z_n} a_n z_1 - f_{n-1}^{[2]}(z) + \beta_{n-1}^{[2]}(z_1)\end{aligned}$$

and for the last row we obtain:

$$0 = \sum_{i=1}^{n-1} \frac{\partial \Phi_n^{[2]}}{\partial z_i} (a_i z_1 + z_{i+1}) + \frac{\partial \Phi_n^{[2]}}{\partial z_n} a_n z_1 - f_n^{[2]}(z) + \beta_n^{[2]}(z_1)$$

The $(n-1)$ first equations give the value of $\Phi^{[2]}(z)$, which cancels all quadratic terms in the $(n-1)$ first lines of $f^{[2]}(z)$. Moreover, as $\beta^{[2]}(z_1)$ is a free homogeneous vector field, it is also possible to cancel some terms of $f_n^{[2]}(z)$. More precisely, setting $\beta_i^{[2]}(z_1) = \beta_{1,i} z_1^2$, for the first equation we have:

$$\Phi_2^{[2]}(z) = -f_1^{[2]}(z) + \beta_{1,1} z_1^2$$

and for the second equation we obtain:

$$\begin{aligned}\Phi_3^{[2]}(z) &= \sum_{i=1}^{n-1} \frac{\partial \left(-f_1^{[2]}(z) + \beta_{1,1} z_1^2 \right)}{\partial z_i} (a_i z_1 + z_{i+1}) \\ &\quad - \frac{\partial f_1^{[2]}(z)}{\partial z_n} a_n z_1 - f_2^{[2]}(z) + \beta_{1,2} z_1^2\end{aligned}$$

thus, we rewrite $\Phi_3^{[2]}(z)$ as follows:

$$\Phi_3^{[2]}(z) = 2\beta_{1,1} z_1 z_2 + \beta_{1,2} z_1^2 + \bar{\Phi}_3^{[2]}(z, \beta_{1,1} z_1^2)$$

where

$$\begin{aligned}\bar{\Phi}_3^{[2]}(z, \beta_{1,1} z_1^2) &= 2\beta_{1,1} a_1 z_1^2 - f_2^{[2]}(z) - \sum_{i=1}^{n-1} \frac{\partial f_1^{[2]}(z)}{\partial z_i} (a_i z_1 + z_{i+1}) \\ &\quad - \frac{\partial f_1^{[2]}(z)}{\partial z_n} a_n z_1\end{aligned}$$

Consequently, the third equation becomes:

$$\begin{aligned}\Phi_4^{[2]}(z) &= \sum_{i=1}^{n-1} \frac{\partial \left(2\beta_{1,1} z_1 z_2 + \beta_{1,2} z_1^2 + \bar{\Phi}_3^{[2]}(z, \beta_{1,1} z_1^2) \right)}{\partial z_i} (a_i z_1 + z_{i+1}) \\ &\quad + \frac{\partial \bar{\Phi}_3^{[2]}(z, \beta_{1,1} z_1^2)}{\partial z_n} a_n z_1 - f_3^{[2]}(z) + \beta_{1,3} z_1^2\end{aligned}$$

And we can rewrite $\Phi_4^{[2]}(z)$ as follows:

$$\Phi_4^{[2]}(z) = 2\beta_{1,1} z_1 z_3 + 2\beta_{1,2} z_1 z_2 + \beta_{1,3} z_1^2 + \bar{\Phi}_4^{[2]}(z, \beta_{1,1} z_1^2, \beta_{1,1} z_1 z_2, \beta_{1,2} z_1^2)$$

where only terms of the form $\beta_{1,i} z_1 z_j$, for $j + i \geq 4$ are considered outside the function $\bar{\Phi}_4^{[2]}$, thus we have:

$$\begin{aligned}\bar{\Phi}_4^{[2]}(z, \beta_{1,1} z_1^2, \beta_{1,1} z_1 z_2, \beta_{1,2} z_1^2) &= 2(\beta_{1,2} a_1 + \beta_{1,1} a_2) z_1^2 + 2\beta_{1,1} z_2^2 + 2\beta_{1,1} a_1 z_1 z_2 \\ &\quad + \sum_{i=1}^{n-1} \frac{\partial \bar{\Phi}_3^{[2]}(z, \beta_{1,1} z_1^2)}{\partial z_i} (a_i z_1 + z_{i+1}) + \frac{\partial \bar{\Phi}_3^{[2]}(z, \beta_{1,1} z_1^2)}{\partial z_n} a_n z_1 - f_3^{[2]}(z)\end{aligned}$$

Recursively, we obtain:

$$\begin{aligned}\Phi_n^{[2]}(z) &= 2z_1 \sum_{i=2}^{n-1} \beta_{1,n-i} z_i + \beta_{1,n-1} z_1^2 \\ &\quad + \bar{\Phi}_n^{[2]} \left(z, \beta_{1,1} \sum_{j \geq i=1}^{n-2} z_i z_j, \beta_{1,2} \sum_{j \geq i=1}^{n-3} z_i z_j, \dots, \beta_{1,n-1} z_1^2 \right)\end{aligned}$$

And finally the last equation gives:

$$0 = \sum_{i=1}^{n-1} (a_i z_1 + z_{i+1}) \frac{\partial \left(2z_1 \sum_{j=2}^{n-1} \beta_{1,n-j} z_j + \beta_{1,n-1} z_1^2 + \bar{\Phi}_n^{[2]}(\dots) \right)}{\partial z_i} \\ + \frac{\partial \bar{\Phi}_n^{[2]}(\dots)}{\partial z_n} a_n z_1 - f_n^{[2]} + \beta_{1,n} z_1^2$$

which gives:

$$-2z_1 \sum_{i=2}^{n-1} \beta_{1,n-i} z_i - \beta_{1,n} z_1^2 = \sum_{i=1}^{n-1} \frac{\partial \bar{\Phi}_n^{[2]}(\dots)}{\partial z_i} (a_i z_1 + z_{i+1}) + \frac{\partial \bar{\Phi}_n^{[2]}(\dots)}{\partial z_n} a_n z_1 - f_n^{[2]} \\ + 2 \sum_{i=1}^{n-1} \beta_{1,n-i} z_i (a_1 z_1 + z_2) + 2\beta_{1,n-1} z_1 (a_1 z_1 + z_2)$$

Consequently, the free vector field $\beta^{[2]}(z_1)$ can only cancel the quadratic term $z_1 z_i$ for all $i \in \{1, \dots, n\}$ in the last equation.

For the second homological equation, we have only $\gamma^{[1]}(z_1)$ as a free vector field. Thus in

$$\gamma^{[1]}(z_1) = -\frac{\partial \Phi^{[2]}}{\partial z} B + g^{[1]}(z)$$

the vector field $\gamma^{[1]}(z_1)$ is only able to cancel terms in z_1 .

References

1. J.-P. Barbot, I. Belmouhoub, and L. Boutat-Baddas, Observability normal forms, LNCIS 295, in *New Trends in Nonlinear Dynamics and Control*, W. Kang et al., Eds., Springer-Verlag, Berlin, Heidenberg, 2003, pp. 3–17.
2. J.-P. Barbot, T. Boukhobza, and M. Djemai, Sliding mode observer for triangular input form, in *Proceedings of the 35th CDC*, Kobe, Japan, 1996.
3. J.-P. Barbot, S. Monaco, and D. Normand-Cyrot, Quadratic forms and approximate feedback linearization in discrete time, *Int. Jour. Control*, 67 (4), 567–586, 1997.
4. G. Bartolini, A. Pisano, and E. Usai, First and second derivative estimation by sliding mode technique, *Int. J. Signal processing*, 4, 167–176, 2000.
5. I. Belmouhoub, M. Djemai, and J.-P. Barbot, Cryptography by discrete-time hyperchaotic systems, in *Proceedings of the 43th IEEE CDC 03*, Maui, Hawaii, 2003.

6. I. Belmouhoub, M. Djemai, and J.-P. Barbot, Observability quadratic normal forms for discrete-time systems, *IEEE-TAC*, 50 (7), 1031–1038, 2005.
7. G. Besançon, A viewpoint on observability and observer design for nonlinear system, *Lecture Notes in Control and Information Sciences*, 244, Springer, 1999, pp. 1–22.
8. D. Bestle and M. Zeitz, Canonical form observer design for nonlinear time varying systems, *Int. J. Control*, 38, 429–431, 1983.
9. J. Birk and M. Zietz, Extended Luenberger observers for nonlinear multivariable systems, *Int. J. Control*, 47, 1823–1836, 1988.
10. L. Boutat-Baddas, D. Boutat, J.-P. Barbot, and R. Tauleigne, Quadratic Observability normal form, in *Proceedings of the 41st IEEE CDC 01*, Orlando, 2001.
11. L. Boutat-Baddas, J.-P. Barbot, D. Boutat, and R. Tauleigne, Observability bifurcation versus observing bifurcations, in *Proceedings of the IFAC World Congress*, Barcelona, 2002.
12. D. Boutat and J.-P. Barbot, Poincaré normal form for a class of driftless systems in a one-dimensional submanifold neighborhood, *Math. Control, Signals Systems*, 15, 256–274, 2002.
13. G. Chen and X. Dong, From chaos to order: methodologies, perspectives and applications, *World Scientific Series on Nonlinear Science, Series A*, vol. 24, World Scientific, 1998.
14. S. Diop, J.W. Grizzle, P.E. Morral, and A.G. Stefanoupolou, Interpolation and numerical differentiation for observer design, in *Proceedings of the American Control Conference*, American Automatic Control Council, Evanston, IL, 1994, pp. 1329–1335.
15. S. Drakunov and V. Utkin, Sliding mode observer: Tutorial, in *Proceedings of the IEEE CDC*, New Orleans, 1995.
16. J.-P. Gauthier and G. Bornard, Observability for any $u(t)$ of a class of bilinear systems, *IEEE Trans. Autom. Control*, 26, 922–926, 1981.
17. J.-P. Gauthier, H. Hammouri, and S. Othman, A simple observer for nonlinear systems: application to bioreactors, *IEEE Trans. Autom. Control*, 37, 875–880, 1992.
18. G. Gu, A. Sparks, and W. Kang, Bifurcation analysis and control for model via the projection method, in *Proceedings of the ACC*, IEEE, Philadelphia, 3617–3621, 1998.
19. B. Hamzi, J.-P. Barbot, S. Monaco, and D. Normand-Cyrot, Nonlinear discrete-time control of systems with a Naimar-Sacker bifurcation, *Syst. Control Lett.*, 44, 245–258, 2001.
20. B. Hamzi, W. Kang, and J.-P. Barbot, On the control of bifurcations, in *Proceedings of the 39th IEEE CDC*, Phoenix, 1999.
21. J. Hauser and Z. Xu, An approximate Frobenius theorem, in *Proceedings of the IFAC World Congress*, IFAC, Sydney, vol. 8, pp. 43–46, 1993.
22. R. Hermann and A.J. Krener, Nonlinear controllability and observability, *IEEE Trans. Autom. Control*, 22, 728–740, 1977.
23. M. Hou and P.C. Muller, Design of observers for linear systems with unknown inputs, *IEEE Trans. Autom. Control*, 37, 871–875, 1992.
24. M. Hou, K. Busawon, and M. Saif, Observer design based on triangular form generated by injective map, *IEEE Trans. Autom. Control*, 45, 1350–1355, 2000.
25. A. Isidori, *Non-linear Control Systems*, 2nd ed., Springer-Verlag, New York, 1989.

26. M. Itoh and L.O. Chua, Reconstruction and synchronization of hyperchaotic circuit via one state variable, *Int. J. Bif. Chaos*, 12 (10), 2069–2085, 2002.
27. W. Kang, Bifurcation and normal form of nonlinear control system: Part I and II, *SIAM J. Control Optim.*, 36, 193–232, 1998.
28. W. Kang and A.J. Krener, Extended quadratic controller normal form and dynamic state feedback linearization of non linear systems, *SIAM J. Control Optim.*, 30 (6), 1319–1337, 1992.
29. W. Kang and A.J. Krener, Nonlinear observer design, a backstepping approach, personal communication.
30. H. Keller, Nonlinear observer design by transformation into a generalized observer canonical form, *Int. J. Control*, 46, 1915–1930, 1987.
31. H.K. Khalil, *Nonlinear System*, Macmillan, New York, 2nd ed., 1995.
32. H.K. Khalil, High-gain observers in nonlinear feedback control, *Lecture Notes in Control and Information Sciences*, vol. 244, Springer, 1999, pp. 249–268.
33. H.J. Korsch and H.-J. Jodl, *Chaos. A Program Collection for the PC*, Springer, 2nd ed., 1998.
34. A.J. Krener, Approximate linearization by state feedback and coordinate change, *Syst. Control Lett.*, 5, 181–185, 1983.
35. A.J. Krener, Feedback linearization mathematical control theory, in *Mathematics Control and Theory*, J. Bailleul and J.-C. Willems (eds), Springer, 1998, pp. 66–98.
36. A.J. Krener and L. Li, Normal forms and bifurcations of discrete time nonlinear control systems, *SIAM J. Control Optim.*, 40, 1697–1723, 2002.
37. A. Krener and A. Isidori, Linearization by output injection and nonlinear observer, *Syst. Control Lett.*, 3, 47–52, 1983.
38. A.J. Krener and W. Respondek, Nonlinear observer with linearizable error dynamics, *SIAM J. Control and Optim.*, 23, 197–216, 1985.
39. A.J. Krener and M.Q. Xiao, Observers for linearly unobservable nonlinear systems, *Syst. Control Lett.*, 46, 281–288, 2002.
40. J.P. LaSalle, Some extentions of Lyapunov's second method, *IRE Trans. Circ. Theory*, CT-7, 520–527, 1960.
41. C. Lobry, Contôlabilité des systèmes non linéaires, *SIAM J. Control.*, 573–605, 1970.
42. E.E. Macau, C. Grebogi, and Y.-C. Lay, Active synchronization in nonhyperbolic hyperchaotic systems, *Phys. Rev. E*, 65, 027202, 2002.
43. H. Nijmeijer and I.M.Y. Mareels, An observer looks at synchronization, *IEEE Trans. Circuits Syst. 1. Fundame. Theory Appl.*, 44 (11), 882–891, 1997.
44. H. Nijmeijer and van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, Berlin, 1990.
45. W. Perruquetti and J.-P. Barbot, *Sliding Mode Control in Engineering*, Marcel Dekker, 2002.
46. F. Plestan and A. Glumineau, Linearization by generalized input-output injection, *Syst. Control Lett.*, 31, 115–128, 1997.
47. H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste*, Gauthier Villard, 1899 Réédition 1987, Bibliothèque Scientifique A. Blanchard.
48. J.H. Peng, E.J. Ding, M. Ding, and W. Yang, Synchronization hyperchaos with a scalar transmitted signal, *Phys. Rev.*, 76 (6), 904–907, 1996.
49. K. Pyragas, Predictable chaos in slightly perturbed unpredictable chaotic systems, *Phys. R. Lett. A*, 181, 203–210, 1993.

50. A. Robertsson and R. Johansson, Observer backstepping for a class of nonminimum-phase systems, in *Proceedings of the 38th IEEE-CDC*, Phoenix, 1999.
51. O.E. Rossler, An equation for hyperchaos, *Phys. Lett.*, 71A (2,3), 1979.
52. J. Rudolph and M. Zeitz, A Block triangular nonlinear observer normal form, *Syst. Control Lett.*, 23, 1–8, 1994.
53. H.J. Sussmann, Single input observability of continuous time systems, *Math. Syst. Theory*, 12, 371–393, 1979.
54. J.-J. Slotine, J.K. Hedrick, and E.A. Misawa, On sliding observer for nonlinear systems, *ASME JDSMC*, 109, 245–252, 1987.
55. I.A. Tall and W. Respondek, Normal forms and invariants of nonlinear single-input systems with noncontrollable linearization, in *Proceedings of the IFAC NOLCOS*, Barcelona, 2001.
56. R. Tauleigne, L. Boutat-Baddas, and J.-P. Barbot, Syntonisation Chaotique, in *Proceedings Journée d'Etude Automatique et Electronique*, Angoulême, France, 12 2002.
57. J. Tsinias, Observer design for nonlinear systems, *Syst. Control Lett.*, 13, 135–142, 1989.
58. M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed., Classics in Applied Mathematics, vol. 42, SIAM, 2002.
59. S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Text in Applied Mathematics, vol. 2, Springer-Verlag, New York, 1998.
60. X. Xia and W. Gao, Nonlinear observer design by observer error linearization, *SIAM J. Control Optim.*, 27, 199–213, 1989.
61. Y. Xiong and M. Saif, Sliding mode observer for nonlinear uncertain systems, *IEEE Trans. Autom. Control*, 46, 2012–2017, 2001.

10

Nonlinear Observer Design for Smooth Systems

A.J. Krener and M. Xiao

CONTENTS

10.1 Introduction	411
10.2 Main Results	415
10.3 Conclusion	421
References	421

10.1 Introduction

Recently, Kazantzis–Kravaris and Kreisselmeier–Engel have suggested two apparently different approaches for constructing observers for nonlinear systems. We show that these approaches are closely related, leading to observers with linear error dynamics in transformed variables. In particular, we give the sufficient conditions for the existence of smooth solutions to the Kazantzis–Kravaris partial differential equation (KK PDE). These methods can be used for systems that exhibit chaotic behavior.

We consider the problem of constructing an observer for a smooth system without controls

$$\begin{aligned}\dot{x} &= f(x) = Fx + \tilde{f}(x) \\ y &= h(x) = Hx + \tilde{h}(x) \\ x(0) &= x^0\end{aligned}\tag{10.1}$$

where $f : \mathcal{X} \rightarrow \mathbb{R}^n$ and $h : \mathcal{X} \rightarrow \mathbb{R}^p$ are C^r functions with $r \geq 1$ and $\tilde{f}(x) = o(x)$, $\tilde{h}(x) = o(x)$. The set $\mathcal{X} \subset \mathbb{R}^n$ is assumed to be an invariant subset of

the dynamics and a neighborhood of the equilibrium $x = 0$. We let $\mathcal{Y} = h(\mathcal{X}) \subset \mathbb{R}^p$. Typically $p \leq n$.

An observer is a second dynamical system

$$\begin{aligned}\dot{\hat{z}} &= a(\hat{z}, y) = A\hat{z} + By + \bar{a}(\hat{z}, y) \\ \hat{x} &= c(\hat{z}, y) = C\hat{z} + Dy + \bar{c}(\hat{z}, y) \\ \hat{z}(0) &= \hat{z}^0\end{aligned}\tag{10.2}$$

where $a : \mathcal{Z} \times \mathcal{Y} \rightarrow \mathbb{R}^k$ and $c : \mathcal{Z} \times \mathcal{Y} \rightarrow \mathbb{R}^n$ are C^r functions, $\mathcal{Z} \subset \mathbb{R}^k$ and $\bar{a}(\hat{z}, y) = o(\hat{z}, y)$, $\bar{c}(\hat{z}, y) = o(\hat{z}, y)$.

The goal is to choose the observer in such a way that the estimation error $\tilde{x}(t) = x(t) - \hat{x}(t) \rightarrow 0$ as $t \rightarrow \infty$. The dimension k of the observer can be different from the dimension n of the system. For nonlinear systems one expects that $k \geq n$. There is a vast literature on this topic; we refer the reader to a recent survey paper [3] and conference proceedings [6].

Kazantzis and Kravaris [1] have introduced a method for constructing an observer which has linear error dynamics in transformed coordinates. We briefly review their method. Suppose the system (10.1) is real analytic, one selects an $n \times p$ matrix B and an invertible $n \times n$ matrix T such that the matrix

$$A = (TF - BH)T^{-1}$$

is Hurwitz and such that the eigenvalues of A are distinct from those of F . Then one seeks a real analytic solution of the KK PDE

$$\frac{\partial \theta}{\partial x}(x)f(x) = A\theta(x) + \beta(h(x))\tag{10.3}$$

where $\beta : \mathcal{Y} \rightarrow \mathbb{R}^n$ is real analytic and

$$\frac{\partial \beta}{\partial y}(y) = B$$

If θ satisfies this PDE then

$$\frac{\partial \theta}{\partial x}(0) = T$$

and so θ is a local diffeomorphism. If we define a change of coordinates

$$z = \theta(x)$$

then

$$\begin{aligned}\dot{z} &= Az + \beta(y) \\ y &= h(\theta^{-1}(z))\end{aligned}\tag{10.4}$$

One can construct a local observer for the transformed system (10.4),

$$\begin{aligned}\dot{\hat{z}} &= A\hat{z} + \beta(y) \\ \hat{x} &= \theta^{-1}(\hat{z}) \\ \hat{z}(0) &= 0\end{aligned}\tag{10.5}$$

which has linear error dynamics in the transformed coordinates

$$\dot{\tilde{z}} = A\tilde{z}\tag{10.6}$$

where $\tilde{z}(t) = z(t) - \hat{z}(t)$. Since A is Hurwitz, the error goes to zero as $t \rightarrow \infty$, provided $x(t)$ stays sufficiently small.

The observer can also be implemented in the original coordinates,

$$\dot{\hat{x}} = f(\hat{x}) + \left(\frac{\partial \theta}{\partial x}(\hat{x}) \right)^{-1} (\beta(y) - \beta(h(\hat{x})))\tag{10.7}$$

This is a standard form for an observer, a copy of this dynamics driven by a gain times the estimation error of some function of y . In this case, the gain varies with $\hat{x}$.

Kazantzis and Kravaris [1] presented sufficient conditions for the solvability of (10.3). We need some definitions to state them. Let $\lambda = (\lambda_1, \dots, \lambda_n)$ denote the spectrum of F . We say that a complex number μ_j is resonant of degree $d > 0$ with the spectrum of F if there is a tuple $m = (m_1, \dots, m_n)$ of non-negative integers

$$\sum_{i=1}^n m_i \lambda_i = \mu_j, \quad \sum_{i=1}^n m_i = d$$

The spectrum of F is in the *Siegel domain* if 0 is in the convex hull of $\lambda_1, \dots, \lambda_n$ in $\mathbb{C}$. The spectrum of F is in the *Poincaré domain* if 0 is not in the convex hull of $\lambda_1, \dots, \lambda_n$. Since F is real, the spectrum of F is in the Poincaré domain iff it is in the open left half plane of $\mathbb{C}$ or it is in the open right half plane of $\mathbb{C}$.

Let $\mu = (\mu_1, \dots, \mu_n)$ denote the spectrum of A . Kazantzis and Kravaris [1] showed that if no μ_j is resonant of any degree d with the spectrum of F and if the spectrum of F is in the Poincaré domain, then for a given real analytic $\beta(y)$ the KK PDE has a unique real analytic solution defined in some neighborhood of $x = 0$.

Suppose $C > 0$, $\nu > 0$. A complex number μ_j is *type* (C, ν) with respect to the spectrum of F if for any tuple $m = (m_1, \dots, m_n)$ of non-negative integers

with $\sum m_i = d > 0$, one has

$$\left| \sum_{i=1}^n m_i \lambda_i - \mu_j \right| \geq \frac{C}{d^v}$$

Recently, we claimed that if the spectrum of A is of type (C, v) with respect to the spectrum of F , then for given real analytic $\beta(y)$ the KK PDE has a unique real analytic solution defined in some neighborhood of $x = 0$ but we have found an error in our proof [4]. One needs the additional assumption that the spectrum of F is of type (C, v) with respect to the spectrum of F [4].

Another approach to observer design has recently been presented by Kreisselmeier and Engel [2]. The purpose of our paper is to show that their approach is closely related to that of Kazantzis and Kravaris and the two approaches taken together yield the existence of smooth solutions to the KK PDE under suitable conditions. Given the system (10.1), Kreisselmeier and Engel construct an observer as follows.

Choose a $k \times k$ Hurwitz matrix A and a $k \times p$ matrix B where $k \geq n$. Given x^0 , let $x(s, x^0)$, $y(s, x^0)$ denote the corresponding state and output trajectories of (10.1). Define

$$z^0 = \theta(x^0) = \int_{-\infty}^0 e^{-As} B y(s, x^0) ds, \quad (10.8)$$

one has to impose suitable conditions so that the integral exists.

If one can find a mapping $x^0 = \psi(z^0)$ which is a left inverse of θ ,

$$\psi(\theta(x^0)) = x^0$$

then one can construct the Kreisselmeier and Engel observer

$$\begin{aligned} \dot{\hat{z}} &= A\hat{z} + By \\ \hat{x} &= \psi(\hat{z}) \end{aligned} \quad (10.9)$$

Kreisselmeier and Engel showed that for a suitable choice of A, B , the mapping θ is injective so that ψ exists. We shall show that the error in the transformed variables $\tilde{z}(t) = z(t) - \hat{z}(t)$ has linear error dynamics (10.6).

Let us review the differences in the approaches of Kazantzis–Kravaris and Kreisselmeier–Engel. The Kazantzis–Kravaris construction applies to real analytic systems and defines an observer of the same dimension as the system. The observer is constructed by a real analytic change of coordinates found by solving the KK PDE. The KK PDE is locally solvable if the spectrum of the linear part of the dynamics is in the Poincaré domain and if the spectrum of the linear part of the observer is not resonant with

that of the dynamics. Our recent work [4], establishes that it is also locally solvable if the spectrum of both F and A are of type (C, ν) with respect to the spectrum of F . This local solution leads to a local observer. Conditions for the global solvability of the KK PDE are not known. The Taylor series of the solution of the KK PDE can be found to any degree by solving a sequence of linear equations for the coefficients and therefore, the Taylor series of the inverse can be found up to the same degree. This is essential if the observer is to be implemented in the transformed coordinates. The inverse change of coordinates need not be found if the observer is implemented in the original coordinates, but the Jacobian of the change of coordinates must be inverted. If the system (10.1) is only C^r and if the spectrum of A is not resonant up to degree $d \leq r$ with the spectrum of F , an approximate solution to the KK PDE, polynomial of degree d , can be found and used to construct a local observer with error dynamics linear to degree d in the transformed coordinates.

The Kreisselmeier–Engel construction applies to Lipschitz continuous systems (10.1) and defines an observer whose dimension is at least as large as that of the system. The observer is constructed by a change of variables found through an integral equation but the change of variables need not be a change of coordinates. It is not guaranteed to be smooth even if the system is. It can be hard to compute explicitly. Its existence depends on growth conditions for the output of the system in negative time. A left inverse of the change of variables must be found. The observer is implemented in the transformed variables where it has linear error dynamics. The Kreisselmeier–Engel observer requires a choice of A and B to define θ by (10.8). When the system (10.1) is C^1 then the convergence of this integral implies that the spectra of A and F are disjoint. We shall show that if (10.3) does not hold for some invertible T then θ is not differentiable.

10.2 Main Results

In this section, we make the following assumptions about the system

$$\begin{aligned}\dot{x} &= f(x) \\ y &= h(x) \\ x(0) &= x^0\end{aligned}\tag{10.10}$$

1. $f : \mathcal{X} \rightarrow \mathbb{R}^n$ and $h : \mathcal{X} \rightarrow \mathbb{R}^p$ are Lipschitz continuous on an invariant set $\mathcal{X} \in \mathbb{R}^n$ with Lipschitz constants L_f and L_h , respectively

2. A is $k \times mk$ Hurwitz matrix and there are $M > 0, a > L_f$ such that

$$|e^{At}| \leq Me^{-at}$$

3. $\beta : \mathcal{Y} \rightarrow \mathbb{R}^k$ and Lipschitz continuous with Lipschitz constant L_β
 4. When the integral exists, θ is defined by

$$z^0 = \theta(x^0) = \int_{-\infty}^0 e^{-As} \beta(y(s, x^0)) \, ds \quad (10.11)$$

THEOREM 1

Under Assumptions 1–4, the map $\theta : \mathcal{X} \rightarrow \mathbb{R}^n$ exists and is Lipschitz continuous.

PROOF Given x^0 , let $x(s, x^0), y(s, x^0)$ denote the corresponding state and output trajectories of (10.1). Now

$$\begin{aligned} x(t, x^0) &= x^0 + \int_0^t f(x(s, x^0)) \, ds \\ |x(t, x^0)| &\leq |x^0| + \int_0^t |f(x(s, x^0))| \, ds \\ |x(t, x^0)| &\leq |x^0| + \int_0^t L_f |x(s, x^0)| \, ds \end{aligned}$$

so by Gronwall's inequality

$$|x(t, x^0)| \leq |x^0| e^{L_f |t|}$$

Hence,

$$\begin{aligned} |y(t, x^0)| &\leq L_h |x^0| e^{L_f |t|} \\ |\beta(y(t, x^0))| &\leq L_\beta L_h |x^0| e^{L_f |t|} \end{aligned}$$

Since, $a - L_f > 0$, the integral exists

$$\begin{aligned} |\theta(x^0)| &\leq \int_{-\infty}^0 \left| e^{-As} \beta(y(s, x^0)) \right| \, ds \\ &\leq \int_{-\infty}^0 M L_\beta L_h |x^0| e^{(a-L_f)t} \, ds \\ &\leq \frac{M L_\beta L_h |x^0|}{a - L_f} \end{aligned}$$

Next we show that θ is Lipschitz continuous. Given two initial conditions $x^0, x^1 \in \mathcal{X}$, with corresponding state and output trajectories $x(t, x^i), y(t, x^i)$, the Lipschitz assumptions imply that

$$\begin{aligned} |x(t, x^1) - x(t, x^0)| &\leq |x^1 - x^0| e^{L_f |t|} \\ |y(t, x^1) - y(t, x^0)| &\leq L_h |x^1 - x^0| e^{L_f |t|} \\ |\beta(y(t, x^1)) - \beta(y(t, x^0))| &\leq L_\beta L_h |x^1 - x^0| e^{L_f |t|} \end{aligned}$$

so

$$\begin{aligned} |\theta(x^1) - \theta(x^0)| &\leq \int_{-\infty}^0 \left| e^{-As} \left(\beta(y(t, x^1)) - \beta(y(t, x^0)) \right) \right| ds \\ &\leq \int_{-\infty}^0 \left| e^{-As} L_\beta L_h |x^1 - x^0| e^{L_f |s|} \right| ds \\ &\leq |x^1 - x^0| \int_{-\infty}^0 M L_\beta L_h e^{(a-L_f)s} ds \end{aligned}$$

The map θ is Lipschitz continuous with Lipschitz constant

$$L_\theta = \int_{-\infty}^0 M L_\beta L_h e^{(a-L_f)s} ds = \frac{M L_\beta L_h}{a - L_f}$$

THEOREM 2

Under Assumptions 1–4, let $x(t), y(t)$ be state and output trajectories of the system where $x(0) \in \mathcal{X}$. Let $z(t) = \theta(x(t))$ where θ is defined by (10.8). Then

$$\frac{d}{dt} z(t) = Az(t) + \beta(y(t))$$

PROOF Because the system is autonomous

$$z(t) = \int_{-\infty}^0 e^{-As} \beta(y(s+t)) ds$$

Let $r = s + t$ then

$$z(t) = e^{At} \int_{-\infty}^t e^{-AR} \beta(y(r)) dr$$

so

$$\dot{z}(t) = Az(t) + \beta(y(t))$$

REMARK 1

It is interesting to note that Theorem 2 does not require any assumption of differentiability.

COROLLARY 1

The Kreisselmeier–Engel observer (10.9) has linear error dynamics in transformed variables.

COROLLARY 2

Under Assumptions 1–4, if $\theta(x)$ is C^1 then $\theta(x)$ satisfies the KK PDE (10.3).

COROLLARY 3

Under Assumptions 1–4, the observer (10.5) has asymptotically stable, linear error dynamics in the z variables.

THEOREM 3

In addition to Assumptions 1–4, assume that f , h , and β are C^1 . Then θ is C^1 .

PROOF Proceeding formally from (10.8) we have

$$\frac{\partial \theta}{\partial x^0}(x^0) = \int_{-\infty}^0 e^{-As} \frac{\partial \beta}{\partial x^0}(y(s, x^0)) ds$$

If this integral converges then it is the actual derivative. By the chain rule

$$\frac{\partial \theta}{\partial x^0}(x^0) = \int_{-\infty}^0 e^{-As} \frac{\partial \beta}{\partial y}(y(s, x^0)) \frac{\partial h}{\partial x}(x(s, x^0)) \frac{\partial x}{\partial x^0}(s, x^0) ds$$

Let

$$\Phi(s, x^0) = \frac{\partial x}{\partial x^0}(s, x^0)$$

$$F(s, x^0) = \frac{\partial f}{\partial x}(x(s, x^0))$$

$$H(s, x^0) = \frac{\partial h}{\partial x}(x(s, x^0))$$

$$B(s, x^0) = \frac{\partial \beta}{\partial y}(y(s, x^0))$$

Now

$$\frac{d}{ds} \Phi(s, x^0) = F(s, x^0) \Phi(s, x^0)$$

$$\Phi(0, x^0) = I$$

so

$$\begin{aligned}\Phi(t, x^0) &= I + \int_0^t F(s, x^0) \Phi(s, x^0) \, ds \\ |\Phi(t, x^0)| &\leq |I| + \int_0^t |F(s, x^0)| |\Phi(s, x^0)| \, ds \\ |\Phi(t, x^0)| &\leq 1 + \int_0^t L_f |\Phi(s, x^0)| \, ds\end{aligned}$$

and by Gronwall's inequality

$$|\Phi(t, x^0)| \leq e^{L_f |t|}$$

Finally

$$\begin{aligned}\int_{-\infty}^0 |e^{-As} B(s, x^0) H(s, x^0) \Phi(s, x^0)| \, ds &\leq \int_{-\infty}^0 M L_\beta L_h e^{(a-L_f)s} \, ds \\ &\leq L_\theta\end{aligned}$$

so $(\partial\theta/\partial x)(x^0)$ exists for $x^0 \in \mathcal{X}$. ■

Suppose f, h are C^1 on $\mathcal{X}$ with Lipschitz continuous derivatives $\partial f/\partial x, \partial h/\partial x$ and β is C^1 on $\mathcal{Y}$ with Lipschitz continuous derivative $\partial\beta/\partial y$. One can show that if a is large enough then $\partial\theta/\partial x$ is Lipschitz continuous. Furthermore, if f, h, β are C^2 then $\partial^2\theta/\partial x^2$ exists. Similar statements hold for the higher derivatives. For C^∞ functions f, h, β and compact $\mathcal{X}$, the larger a the more derivatives of θ that can be shown to exist.

THEOREM 4

In addition to Assumptions 1–4, assume that f, h , and β are C^1 and

$$\begin{aligned}\frac{\partial f}{\partial x}(0) &= F \\ \frac{\partial h}{\partial x}(0) &= H \\ \frac{\partial \beta}{\partial y}(0) &= B\end{aligned}$$

Then

$$\frac{\partial \theta}{\partial x}(0) = T$$

where T is the unique solution of

$$TF - AT = BH \quad (10.12)$$

PROOF If λ is an eigenvalue of F then $|\lambda| \leq L_f$ and if μ is an eigenvalue of A then $|\mu| \geq a$. Since $a > L_f$, the spectra of F and A are disjoint. Therefore, (10.12) has a unique solution T . From the definition of θ , we have

$$\frac{\partial \theta}{\partial x}(0) = \int_{-\infty}^0 e^{-As} B H e^{Fs} ds$$

Let

$$S = \int_{-\infty}^0 e^{-As} B H e^{Fs} ds$$

then

$$\int_{-\infty}^0 \frac{d}{ds} e^{-As} B H e^{Fs} ds = BH$$

$$SF - AS = BH \quad (10.13)$$

Therefore, T and S satisfy Equation (10.12) and Equation (10.13) so $T = S$. ■

COROLLARY 4

In addition to Assumptions 1–4, assume that f , h , and β are C^1 . If F, H, A, B are not related by (10.12) for some T then θ (10.11) is not differentiable at $x = 0$.

COROLLARY 5

In addition to Assumptions 1–4, assume that $k = n$ and f, h , and β are C^1 then θ is a local diffeomorphism iff the unique T satisfying (10.12) is invertible.

THEOREM 5

Suppose the spectra of F and A are disjoint and T satisfies (10.12). If (H, F) is not observable then T is not invertible.

PROOF Suppose (H, F) is not observable then there exist $\lambda \in \sigma(F)$ and a vector $x \in \mathbb{R}^n$, $x \neq 0$ such that $Hx = 0$ and $Fx = \lambda x$. We multiply (10.12) by x to obtain

$$\lambda TX - Tax = 0$$

If T is invertible the $Tx \neq 0$ so λ is an eigenvalue of A which is a contradiction. ■

THEOREM 6

Suppose the spectra of F and A are disjoint and T satisfies (10.12). If (A, B) is not controllable then T is not invertible.

PROOF Suppose (A, B) is not controllable then there exist $\mu \in \sigma(A)$ and a vector $\xi \in \mathbb{R}^n$, $\xi \neq 0$ such that $\xi'B = 0$ and $\xi'A = \mu\xi'$. We multiply (10.12) by ξ' to obtain

$$\xi'TF - \mu\xi'T = 0.$$

If T is invertible the $\xi'T \neq 0$ so μ is an eigenvalue of F which is a contradiction. ■

10.3 Conclusion

We have shown that the approaches of Kazantzis–Kravaris and Kreisselmeier–Engel to observer design are closely related. Both lead to observers with linear error dynamics in transformed variables. The former requires the solution of a PDE and the latter requires multiple solutions to an ODE followed by quadratures. From an implementation point of view, the former is easier as the PDE can be solved approximately by a finite power series; but this solution is only local as is the resulting observer. Neither approach has been generalized to systems with inputs yet.

These methods can be used to construct observers for systems that can exhibit chaotic behavior as was shown earlier [5].

References

1. N. Kazantzis and C. Kravaris, Nonlinear observer design using Lyapunov's auxiliary theorem, *Syst. Control Lett.*, 34, 241–247, 1998.
2. G. Kreisselmeier and R. Engel, Nonlinear observers for autonomous Lipschitz continuous systems, *IEEE Trans. Autom. Contr.*, 48 (3), 451–464, 2003.
3. A.J. Krener, Nonlinear observers, in control of nonlinear systems, in *Encyclopedia of Life Support Systems*, H. Unbehauen, Ed., Eolss Publishers, Oxford, UK, 2004 [<http://www.eolss.net>].

4. A.J. Krener and M. Xiao, Nonlinear observer design in the Siegel domain, *SIAM J. Control Optim.*, 41 (3), 932–953, 2002.
5. A. Krener and M.Q. Xiao, Observer for linearly unobservable nonlinear systems, *Syst. Control Lett.*, 46, 281–288, 2002.
6. H. Nijmeijer and T. Fossen, *New Directions in Nonlinear Observer Design*, Springer-Verlag, New York, 1999.

Part III

Some Applications

11

Chaos and Communications

R. Quéré, J. Guittard, and J.C. Nallatamby

CONTENTS

11.1	Introduction	426
11.2	Synchronization of Chaotic Systems	426
11.2.1	Peccora–Caroll (PC) Synchronization	426
11.2.1.1	Principle of PC Synchronization	426
11.2.1.2	Example	428
11.2.2	Feedback-Type Synchronization	430
11.2.3	Synchronization by the Inverse System Approach	431
11.3	Communications with Chaos	433
11.3.1	Chaos Masking	433
11.3.2	Chaos Modulation	433
11.3.3	Chaos Shift Keying	434
11.3.3.1	Coherent CSK	434
11.3.3.2	Noncoherent CSK	435
11.4	Analysis of a Microwave Chaotic Oscillator	437
11.4.1	Analysis of the Voltage Controlled Oscillator	437
11.4.1.1	Equations of the VCO	438
11.4.1.2	Characteristics of the VCO	441
11.4.2	Equation of the Chaotic Oscillator	441
11.5	Chaotic Modulator and Demodulator	447
	References	451

11.1 Introduction

Techniques for the detection of unstable regimes of microwave nonlinear circuits have been presented in the previous chapters. These techniques are capable of predicting the local and the global stabilities of these circuits, and in some cases are also capable of detecting the bifurcation routes leading

to chaotic behaviors. However, beyond these techniques that are mainly used to avoid the initiation of chaotic oscillations in nonlinear circuits, there has been considerable interest in the design of chaotic circuits for communication purposes.

This interest has been aroused by the demonstration of Peccora and Carroll [1] of the ability to synchronize two chaotic systems. Following this, researchers tried to exploit this synchronization property to transmit information, using chaotic waves. Various techniques have been proposed that require the generation of analog or digital chaotic signals, which still remains a problem especially in the microwave domain.

In this chapter, we will present an overview of the different modulation schemes that allow the transmission of some information with chaotic carriers. The particular problems of designing chaotic generators in the microwave domain will be pointed out. Indeed, special techniques for simulating such complex behaviors have to be used. It will be shown by a simple example how the transient envelope can cope with the simulation of chaotic circuits in the microwave domain.

Section 11.2 will be devoted to the presentation of the various synchronization techniques that can be used. In Section 11.3, chaotic modulations for information transmission are reviewed, and Section 11.4 presents an insight into the simulation technique, using a simple chaotic generator example. Finally, an example of transmission of binary data will be discussed in Section 11.5 using the previous chaotic generator.

11.2 Synchronization of Chaotic Systems

11.2.1 Peccora–Carroll (PC) Synchronization

11.2.1.1 Principle of PC Synchronization

This kind of synchronization is the first that has been reported for chaotic systems. It relies on the decomposition of the initial chaotic into two subsystems. The receiver system is a replica of one of the two subsystems.

Specifically, consider a chaotic autonomous system described by the n -dimensional nonlinear equation:

$$\dot{\bar{x}} = f(\bar{x}) \quad \bar{x} \in R^n \quad (11.1)$$

Splitting (11.1) into two subsystems $\dot{\bar{x}} = [\bar{x}_s, \bar{x}_d]^T$ leads to

$$\begin{aligned} \dot{\bar{x}}_s &= g(\bar{x}_s, \bar{x}_d) \quad \bar{x}_s \in R^s \\ \dot{\bar{x}}_d &= h(\bar{x}_s, \bar{x}_d) \quad \bar{x}_d \in R^d \end{aligned} \quad (11.2)$$

Then, the receiver is constituted of a replica of the second subsystem and the receiver state variable equation reads:

$$\dot{\bar{x}}_r = h(\bar{x}_s, \bar{x}_r) \quad \bar{x}_d \in R^d \quad (11.3)$$

where $\bar{x}_r \in R^d$ is the state vector which has to be recovered by the slave system and $\bar{x}_s \in R^s$ is a common state vector for the master and slave systems. The architecture of the synchronization master–slave system is illustrated in Figure 11.1.

The two systems will synchronize if the difference between the transmitted and received state vectors converges to zero as time goes to infinity:

$$\Delta \bar{x} = \bar{x}_r - \bar{x}_d \xrightarrow[t \rightarrow \infty]{} 0$$

Thus, the first time domain derivative of the difference $\Delta \bar{x}$ has to satisfy the nonlinear differential equation:

$$\Delta \dot{\bar{x}} = \dot{\bar{x}}_r - \dot{\bar{x}}_d = h(\bar{x}_s, \bar{x}_r) - h(\bar{x}_s, x_d) \quad (11.4)$$

Provided the two phase trajectories in the phase space are sufficiently close to each other, the difference $\Delta \bar{x}$ satisfies the linearized equation:

$$\Delta \dot{\bar{x}} = \left[\frac{\partial h(x_s, x_d)}{\partial x_d} \right] \cdot \Delta \bar{x} + o(\Delta \bar{x}^2) \quad (11.5)$$

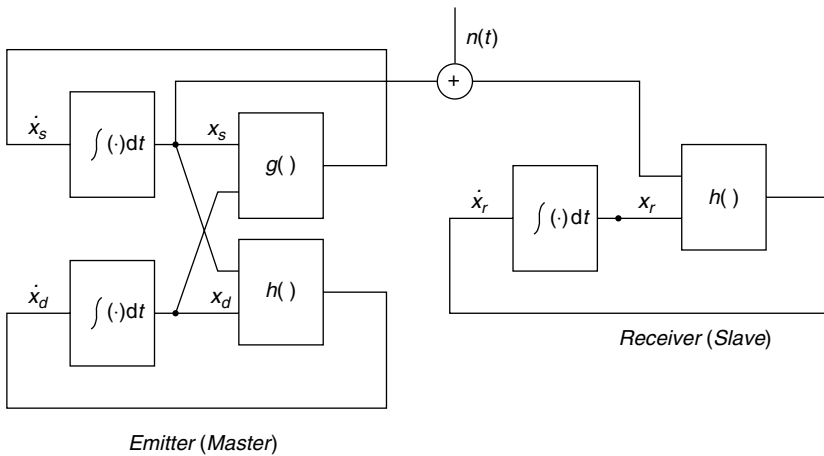


FIGURE 11.1

Architecture of the master–slave synchronization of two chaotic systems.

where $[\partial h(x_s, x_d)/\partial x_d] \in \mathbb{R}^{d \times d}$ is a time variant matrix that will give the stability of the solution of (11.5). However, as the drive of the slave subsystem is a chaotic signal, the stability of the solution must be investigated using the Lyapounov exponent concept [1]. These exponents are calculated from the eigenvalues of a time domain matrix $A(t)$ defined as

$$\begin{aligned}\Delta \dot{A} &= \left[\frac{\partial h(x_s, x_d)}{\partial x_d} \right] \cdot A(t) \\ A(0) &= I\end{aligned}\tag{11.6}$$

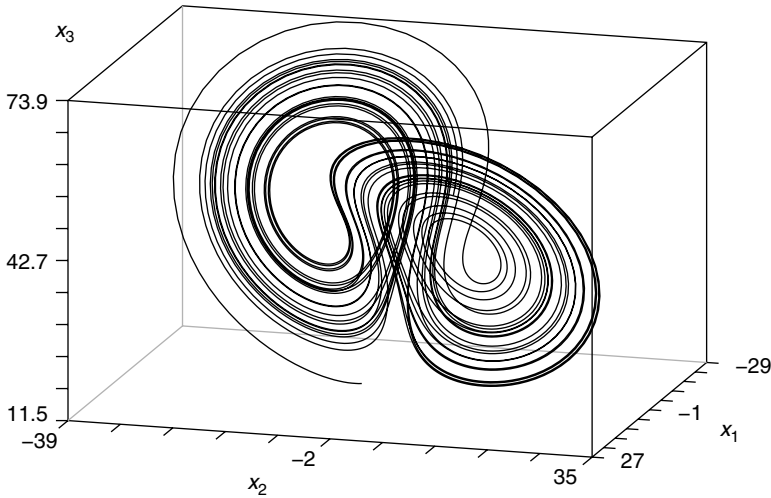
where ΔA is a matrix that is substituted to the vector $\Delta \bar{x}$ and I is the identity matrix. A number of algorithms are available for calculating the Lyapounov exponents from the eigenvalues of $A(t)$ [2]. The solution of (11.5) will be stable if all the Lyapounov exponents are negative. Moreover, the magnitude of these exponents gives an indication of the rate of convergence toward the stable solution $\Delta \bar{x} = 0$.

11.2.1.2 Example

To illustrate the aforementioned point, we will analyze in greater detail the historical example which is based on the Lorenz system. This three-dimensional system is autonomous and reads:

$$\begin{aligned}\dot{x}_1 &= \sigma \cdot (x_2 - x_1) \\ \dot{x}_2 &= -x_1 \cdot x_3 + r \cdot x_1 - x_2 \\ \dot{x}_3 &= x_1 \cdot x_2 - b \cdot x_3\end{aligned}\tag{11.7}$$

where σ , r , and b are three parameters whose values will determine the type of regime that will take place for this set of equations. For example, the phase portrait of a chaotic regime for a particular set of parameters is given in Figure 11.2.

**FIGURE 11.2**

Phase portrait of the Lorenz system in the chaotic regime for $\sigma = 16$, $b = 4$, and $r = 45.92$.

The slave subsystem can be chosen as one of the three couples of state variables, the third one being chosen as the common variable driving the slave subsystem. Thus, the decomposition can take one of the three following possibilities:

$$\bar{x}_d = [x_1, x_2]; x_s = x_3 \quad \text{Case 1}$$

$$\bar{x}_d = [x_1, x_3]; x_s = x_2 \quad \text{Case 2}$$

$$\bar{x}_d = [x_2, x_3]; x_s = x_1 \quad \text{Case 3}$$

In Ref. [1] it is shown that only Case 2 and Case 3 can lead to the synchronization of the slave system. For Case 2, we show in Figure 11.3 the waveforms in two cases corresponding to (a) synchronization and (b) no synchronization. Moreover, in Figure 11.4, the rate of convergence can be determined by plotting the magnitude (in log scale) of the error signal $\Delta \bar{x}$ versus time. From this plot it can be noticed that the rate of convergence is higher for the state variable x_1 than for x_3 .

However, one major disadvantage of this synchronization technique is its sensitivity to a mismatch of parameters between the emitter and the receiver. For example, we have plotted, for the same system, the magnitude of the error vector for two values of the mismatch corresponding 1 and 5% variation of the receiver's parameters. From Figure 11.5 it can be noticed that as the mismatch increases there is a sharp increase in the floor of the magnitude of the error signal. Moreover, in the case of wireless transmission of the drive signal, the received signal at the slave front-end is corrupted by noise.

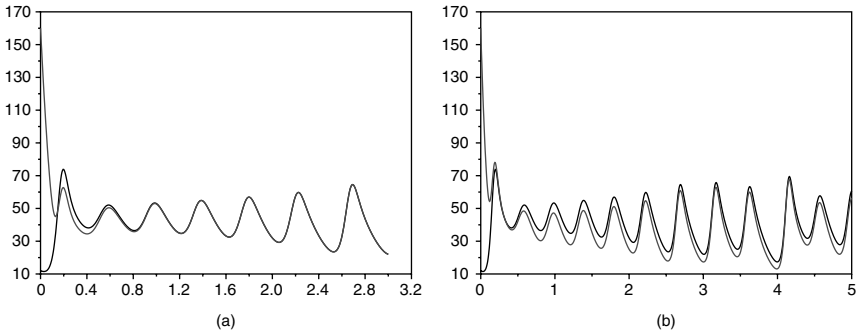


FIGURE 11.3
 x_1 waveforms corresponding to: (a) synchronization and (b) lack of synchronization.

11.2.2 Feedback-Type Synchronization

As seen from the previous example, PC synchronization suffers from a strong sensitivity to parameters mismatch as well as to noisy inputs. The main reason for this sensitivity is that this kind of synchronization is basically an open-loop synchronization [3]. Thus new techniques based on a

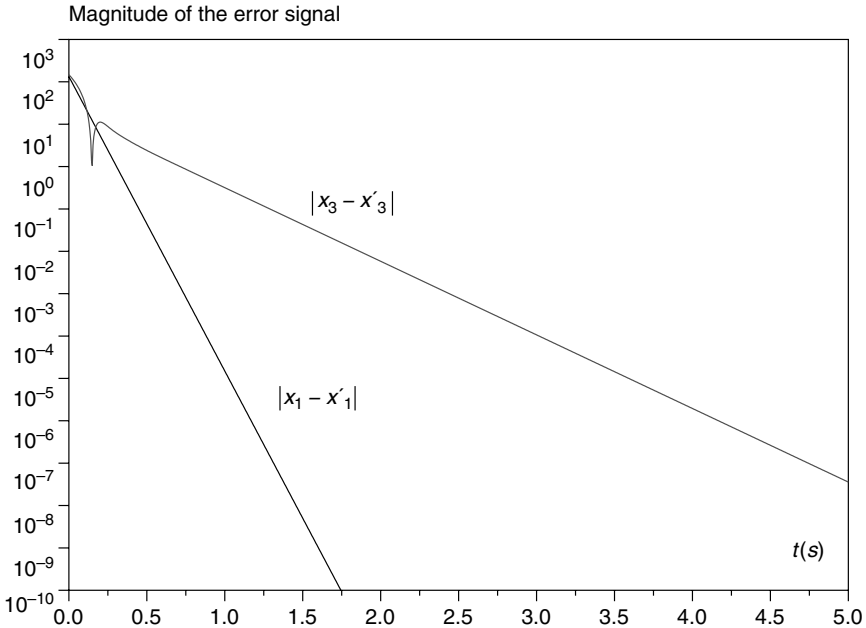
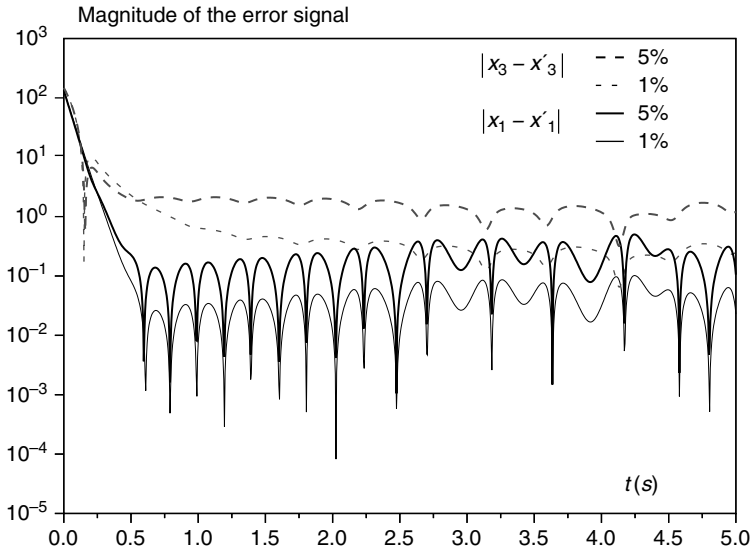


FIGURE 11.4
 Plots of the error signal versus time for a perfect match of the emitter and receiver parameters.

**FIGURE 11.5**

Plots of the magnitude of the error signal in the case of a mismatch of master and slave parameters.

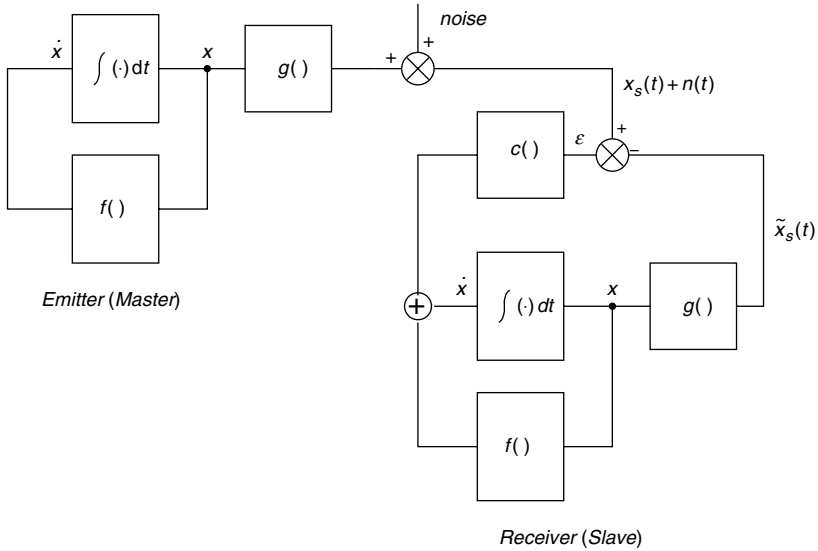
feedback loop have been developed to recover the input signals. These techniques use a replica of the original system and the transmitted signal $x_s(t)$ is compared to a locally regenerated signal $\tilde{x}_s(t)$ to produce an error signal $\varepsilon(t)$ which is fed back into the system through a correction function $c(\cdot)$ as shown in Figure 11.6. Then, the equation of the receiver system reads:

$$\dot{\tilde{x}} = f(\tilde{x}) + c(\varepsilon(t)) = f(\tilde{x}) + c(x_s(t) - \tilde{x}_s(t)) \quad (11.8)$$

By a proper choice of the correction function, it can be shown [4] that under certain conditions the global stability of the steady state condition $\varepsilon(t) = 0$ can be maintained. Moreover, this approach has been recently used to synchronize two different pairs of chaotic oscillators [5].

11.2.3 Synchronization by the Inverse System Approach

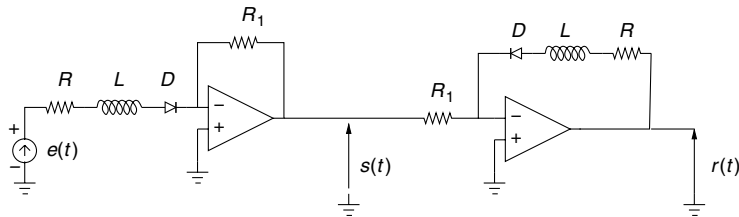
The third type of synchronization system works for non-autonomous systems. A formal definition of such inverse systems is given in Ref. [6]. An example of such a system is given in Figure 11.7 [7]. The diode circuit is well known for producing chaotic signals under certain conditions of

**FIGURE 11.6**

Architecture of a feedback-type chaotic synchronization.

frequency and amplitude of the input generator. The chaotic current is then transformed into an output voltage through the operational amplifier shown in the figure. In contrast, the nonlinear diode resonant circuit of the receiver is fed by the current obtained from the emitter circuit. Thus, to ensure a null current at the input of the second operational amplifier, the output of the receiver $r(t)$ must match the input signal $e(t)$.

When the input signal is a modulated one, the received signal will carry the same modulation information when synchronization holds.

**FIGURE 11.7**

Example of a non-autonomous synchronization system based on an inverse system.

11.3 Communications with Chaos

In Section 11.2, a brief review of chaos synchronization techniques has been given. From this idea of chaos synchronization capability, a number of chaos-based communications systems have been proposed with two goals in mind.

The first one concerns spread spectrum systems where chaotic circuits can be an alternative to more expensive spreading circuits that work in the baseband domain. This can provide resistance against interferences in a multipath environment, but up to now, no chaotic synchronization systems have been used for CDMA transmissions. This is mainly due to the sensitivity of chaotic synchronizers to noise and to the absence of periodical reference in the chaotic spreading signal. A number of chaotic circuits have been used for CDMA systems. Samples of chaotic waves (generally digitally generated) are used instead of usual spreading and coding sequences [8, 9], or an overhead channel is required for transmitting synchronization signals [10].

The second one concerns secure communications and the use of chaotic signals for cryptographic purposes. Most of the works dealing with that subject are performed in the digital domain [11].

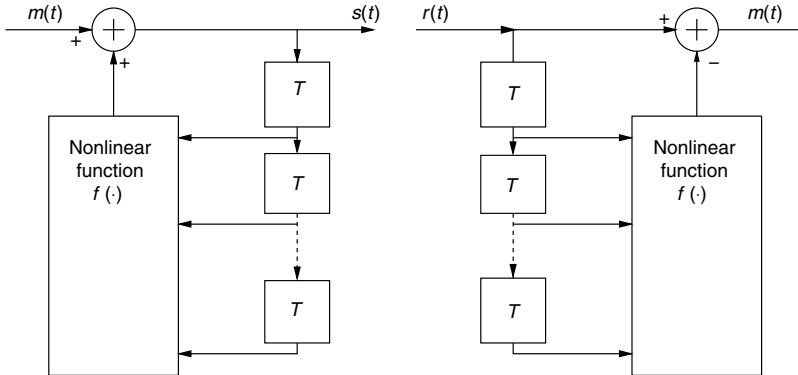
Three main methods for communicating with chaos can be distinguished.

11.3.1 Chaos Masking

In this technique, the information signal is simply added to the chaotic signal in order to mask it. If the level of the information signal is low when compared with the level of the chaotic signal, synchronization remains possible at the receiver front-end. However, in a noisy environment as the masking is an additive process, the information signal cannot be separated from noise at the receiver output. Thus this technique is not very useful for transmission through noisy channels.

11.3.2 Chaos Modulation

In the chaos modulation method, the information signal is directly included in the dynamics of the chaotic generator. This is especially possible for synchronization by inverse systems. For example, for the system in Figure 11.7 the amplitude, frequency, or phase of the input generator can be modulated either in an analog or in a digital way. In this case, after a transient phase of synchronization, the receiver remains synchronized, thus allowing high

**FIGURE 11.8**

Typical coder–decoder based on an inverse chaotic system.

bit rates. Other types of chaos modulation have been proposed either in the low-frequency range, using DSP implementations [12] or in the optical range [13].

A typical coder–decoder system of this type working in the low-frequency range is given in Figure 11.8. A full example of chaos transmission using such a system will be given in Section 11.4.

11.3.3 Chaos Shift Keying

There are various types of chaos shift keying (CSK) that can be used. They can be classified into coherent and noncoherent techniques. All these techniques rely on the association of a digital information with a chaotic wave.

11.3.3.1 Coherent CSK

In a coherent CSK system, the information symbol, say a binary digit, is associated with two chaotic signals $s_0(t)$ and $s_1(t)$ corresponding to the values “0” or “1” as shown in Figure 11.9.

Signals $s_0(t)$ and $s_1(t)$ can be issued either from the same chaotic attractor with different initial conditions or from two different attractors produced by different circuits. The receiver is constituted of two slave circuits, which are replica of the emitter circuits and which have to be synchronized either on $s_0(t)$ or on $s_1(t)$. Decision of the transmitted bit is taken by comparing the outputs of the two correlators shown in Figure 11.9. The main disadvantage of this coding system is that the synchronization is lost at every change of the received bit; therefore, this synchronization has to be recovered during

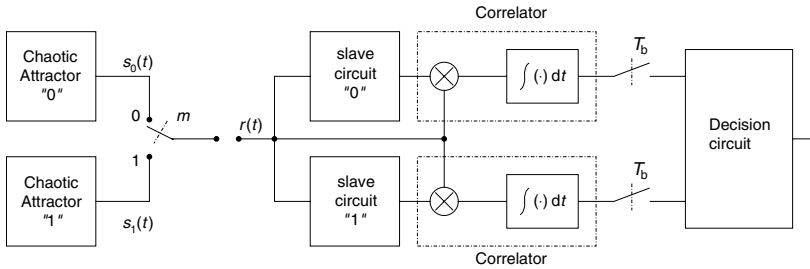


FIGURE 11.9
Example of a coherent CSK system.

each bit duration T_b . However, due to the noisy channel and possible mismatches between master and slave systems, the degree of synchronization is not well controlled, in addition to the synchronization time T_s that has to be greater than T_b . Thus the transmission rate is severely limited by this synchronization time as a loss of synchronization produces catastrophic effects on the bit error rate (BER) of the whole system.

11.3.3.2 Noncoherent CSK

In noncoherent or asynchronous CSK, the information is carried by the statistical properties of the transmitted chaotic wave. The simplest way to do that is to associate the transmission of a chaotic signal to a bit "1," and no transmission for a bit "0." This kind of modulation is name chaotic on–off keying (COOK) and the receiver shown in Figure 11.10 has to measure the variance of the received signal. This type of receiver, measuring the variance of the received signal, is sensitive to the level of noise present at the input of the system. Thus the threshold level of the decision circuit must be adapted to the noise level at the input.

This drawback can be eliminated using a differential chaos shift keying DCSK modulation. In this modulation scheme, the slot of time allocated to a bit T_b is divided into two parts. During the first time slot $[0, T_b/2]$, a reference chaotic signal $s(t)$ is transmitted while during the second slot

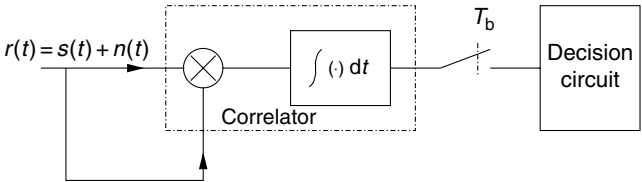
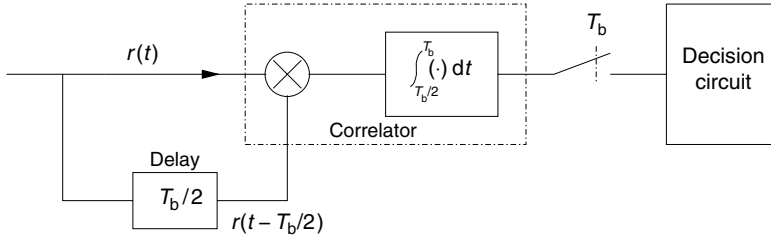


FIGURE 11.10
Architecture of a COOK system.

**FIGURE 11.11**

Architecture of a DCSK system.

$[T_b/2, T_b]$ either $s(t)$ or $-s(t)$ is transmitted depending of the value of the information bit to be transmitted. The receiver system has the structure presented in Figure 11.11.

The output of the correlator at time $t = T_b$ is given by:

$$R = \int_{T_b/2}^{T_b} r(t) \cdot r\left(t - \frac{T_b}{2}\right) dt \quad (11.9)$$

$$R = \int_{T_b/2}^{T_b} ((-1)^m s(t) + n(t)) \cdot \left(s\left(t - \frac{T_b}{2}\right) + n\left(t - \frac{T_b}{2}\right)\right) dt$$

where $n(t)$ is the noise signal introduced by the propagation channel and $m \in \{0, 1\}$ is the information bit to be transmitted. Thus the integral given in (11.8) can be decomposed into four terms [14] as shown in (11.10).

$$R = (-1)^m \int_{T_b/2}^{T_b} s(t) \cdot s\left(t - \frac{T_b}{2}\right) dt + (-1)^m \int_{T_b/2}^{T_b} s(t) \cdot n\left(t - \frac{T_b}{2}\right) dt$$

$$+ \int_{T_b/2}^{T_b} s\left(t - \frac{T_b}{2}\right) \cdot n(t) + \int_{T_b/2}^{T_b} n(t) \cdot n\left(t - \frac{T_b}{2}\right) dt \quad (11.10)$$

The first term in (11.10) is equal to $\pm E_b/2$ while the second and third ones are zero. Only the fourth one depends on the bit duration as its variance increases with integration time. Thus in this system the threshold level for the decision circuit does not depend on the noise level at the input of the receiver. However, one disadvantage of the DCSK modulation method is the need for a time slot for transmission of the reference signal during each bit duration. This slows down the transmission rate.

11.4 Analysis of a Microwave Chaotic Oscillator

As pointed out in the previous section, the capability of transmitting information in a spread spectrum system requires generation of chaotic signals that are well controlled. Thus the design of chaotic oscillators capable of working in the high-frequency range remains a challenge for engineers [15]. Indeed simple design methods as well as efficient predictive CAD tools remain to be developed. This is due to the fact that synchronization of analog chaotic oscillators is very sensitive to parameters and distortions introduced by the transmission channel. Thus the impact of parasitics and real-life nonlinearities that are encountered in practical electronic circuits have to be taken into account. This necessitates simulation of systems or part of the systems at the circuit level.

However, for wireless transmissions, the chaotic signals are band limited around the carrier frequency. This complicates the simulation process, as traditional time domain methods are not suitable because of the very different time scales that are involved. Thus we have to resort to mixed time-frequency methods such as the transient envelope method [16].

In this section we will present an example of the simulation of a particular chaotic oscillator which is well suited for band limited chaotic signal generation in the microwave range. It will be shown how the transient envelope technique allows to derive the bifurcation diagram of the oscillator as well as to predict the performances of a chaotic modulator–demodulator system.

11.4.1 Analysis of the Voltage Controlled Oscillator

The general architecture of the chaotic oscillator is shown in Figure 11.12. It consists of a microwave voltage controlled oscillator (VCO), associated

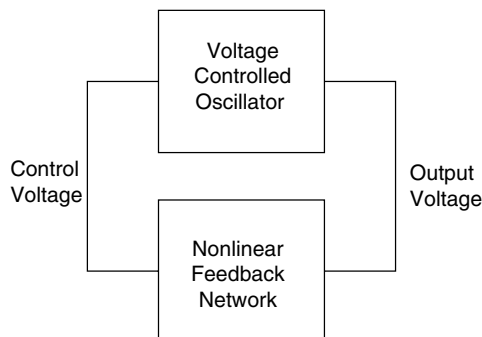


FIGURE 11.12

General architecture of the chaotic oscillator.

with a low-frequency feedback loop. In our case, we will deal with a frequency chaotically modulated oscillator, so the low-frequency loop will involve a frequency discriminator as well as a nonlinear low-frequency filter. Depending of the feedback coefficient, different kinds of regimes can be obtained, ranging from the sinusoidal steady state regime to a hyperchaotic regime.

11.4.1.1 Equations of the VCO

$$\begin{aligned}
 i_{NL}(v_1) &= a \cdot v_1 + b \cdot v_1^3 \quad a = b = 1.04072 \\
 q_1(v_1) &= Q_0 \sqrt{1 + \frac{v_1 + v_r}{\Phi}} \quad Q_0 = 0.7 \text{ pC} \quad \Phi = 0.7 \text{ V} \\
 R &= 5 \Omega \quad L = 3.18 \text{ nH} \quad C_2 = 0.133 \text{ pF} \quad R_2 = 20 \text{ k}\Omega
 \end{aligned} \tag{11.11}$$

The structure of the VCO used in this example is presented in Figure 11.13. It is a simplified Colpitts oscillator where the biasing networks have not been represented for the sake of clarity. A fictitious voltage probe [17] has been added to the circuit, only for simulation purposes. This probe constraints the solution search process to converge to the nontrivial solution. The frequency of the VCO can be adjusted through the nonlinear capacitor control voltage V_r . With the notations shown in Figure 11.13, the

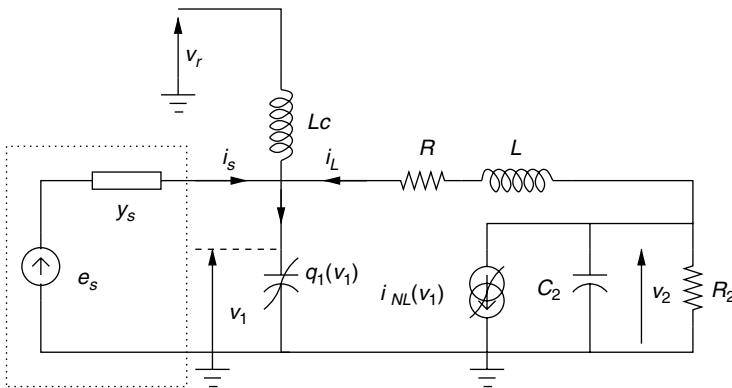


FIGURE 11.13
Schematic of the VCO used for the chaotic oscillator.

network equations of the oscillator are given by:

$$\begin{cases} \frac{dq_1(v_1)}{dt} - i_L - y_s * (e_s(t) - v_1) = 0 \\ C_2 \frac{dv_2}{dt} + i_{NL}(v_1) + \frac{v_2}{R_2} + i_L = 0 \\ L \frac{di_L}{dt} + Ri_L + v_1 - v_2 = 0 \end{cases} \quad (11.12)$$

The constraint equation ensuring that the voltage probe does not perturb the oscillation condition reads:

$$y_s * (e_s(t) - v_1) = 0 \quad (11.13)$$

where $*$ stands for the convolution product.

Define the vectors of variables as:

$$\begin{aligned} \bar{x} &= [v_1, v_2, i_L]^T, \quad \bar{q}(\bar{x}) = [q_1(v_1), C_2 v_2, Li_L]^T, \\ \bar{f}(\bar{x}) &= \left[-i_L, i_{NL} + \frac{v_2}{R_2} + i_L, Ri_L + v_1 - v_2 \right]^T \text{ and} \\ \bar{y}_s &= [y_s, 0, 0]^T; \quad \bar{e}_s = [E \cos(\eta \cdot \omega_0 t), 0, 0]^T \end{aligned}$$

where E and η are, respectively, the amplitude and the reduced frequency of the probe generator which act as two additional variables.

Equations (11.12) can be rewritten under the general form:

$$\frac{d\bar{q}(\bar{x})}{dt} + \bar{f}(\bar{x}) + \bar{y}_s * (\bar{e}_s - v_1) = 0 \quad (11.14)$$

which has to be solved under the constraint (11.12) rewritten in vector form

$$\bar{y}_s * (\bar{e}_s - v_1(\bar{x})) = 0 \quad (11.15)$$

Equations (11.14) is then transformed using the decomposition of the time derivative d/dt into partial derivatives versus two virtual time scales t_c and t_e corresponding to the time scales relative to the carrier and the envelope, respectively [18]:

$$\frac{d}{dt}(\cdot) = \left(\frac{\partial}{\partial t_c} + \frac{\partial}{\partial t_e} \right)(\cdot) \quad (11.16)$$

Thus (11.14) becomes:

$$\frac{\partial \bar{q}(\bar{x}(t_c, t_e))}{\partial t_c} + \frac{\partial \bar{q}(\bar{x}(t_c, t_e))}{\partial t_e} + \bar{f}(\bar{x}(t_c, t_e)) + \bar{y}_s * (\bar{e}_s - v_1) = 0 \quad (11.17)$$

Taking the discrete fourier transform of (11.17) versus the time scale t_c gives a differential equation on the Fourier coefficients of the nonlinear functions as follows:

$$\frac{\partial \bar{Q}(\bar{X}(t_e))}{\partial t_e} + \eta \bar{\Omega}_0 \bar{Q}(\bar{X}(t_e)) + \bar{F}(\bar{X}(t_e)) + \bar{Y}_s \cdot (\bar{E}_s - \bar{X}(t_e)) = 0 \quad (11.18)$$

which is the envelope equation. In this equation, the vector $\bar{X}$ is constituted of the Fourier coefficients of the decomposition of $\bar{x}(t)$ as:

$$\bar{x}(t) = \bar{X}_0 + \sum_{h=1}^H \bar{X}_h^R \cos(h\eta\omega_0 t) - \bar{X}_h^I \sin(h\eta\omega_0 t) \quad (11.19)$$

$\bar{\Omega}_0$ is a block diagonal matrix which reads:

$$\bar{\Omega}_0 = \begin{bmatrix} 0 & 0 & \cdots & & \\ 0 & \begin{bmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{bmatrix} & 0 & & \\ \vdots & & \ddots & & \\ & & & \begin{bmatrix} 0 & -H\omega_0 \\ H\omega_0 & 0 \end{bmatrix} \end{bmatrix} \quad (11.20)$$

$\bar{Y}_s$ is also a diagonal matrix with two nonzero terms:

$$\bar{Y}_s = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & Y_0 & 0 & \cdots & \vdots \\ \vdots & 0 & Y_0 & 0 & 0 \\ & \vdots & 0 & 0 & \cdots \\ & & \vdots & \vdots & \ddots \end{bmatrix} \quad (11.21)$$

Thus system (11.18) is constituted of $3 \times (2H + 1)$ nonlinear equations in $3 \times (2H + 1) + 2$ unknowns $[\bar{X}, E, \eta]$, and the two constraint equations that allow to make the system well conditioned.

At each envelope instant t_e , the system (11.18) is solved using a harmonic balance process organized as a two-tiers procedure. The inner loop

solves (11.18) for fixed valued of $[E, \eta]$ and then these two values are updated in the outer loop to satisfy the constraint equation.

11.4.1.2 Characteristics of the VCO

First the envelope equation is solved in the steady state regime for constant envelope conditions ($\partial/\partial t_e = 0$). This gives the classical harmonic balance conditions for oscillators. The local stability of the solution is checked using the generalized eigenvalues approach. This oscillation was found to be stable for the following value of the resistance $R = 5 \Omega$. Then the frequency versus control voltage of the oscillator was obtained for the various steady state conditions corresponding to $-8 \text{ V} \leq V_r \leq -1 \text{ V}$. This characteristic is shown in Figure 11.14 and the amplitude of the oscillation versus the control voltage in Figure 11.15.

11.4.2 Equation of the Chaotic Oscillator

Using the previous VCO, the chaotic oscillator [19, 20] is built by adding the low-frequency feedback network shown in Figure 11.16. The output $\eta(t)$ of the frequency discriminator is proportional to the reduced frequency $\eta(t)$. The nonlinear filter is constituted by a delay T followed by a nonlinear function that can be realized either in an analog way or digitally with a digital signal processor. This function is the one proposed in Ref. [13] and

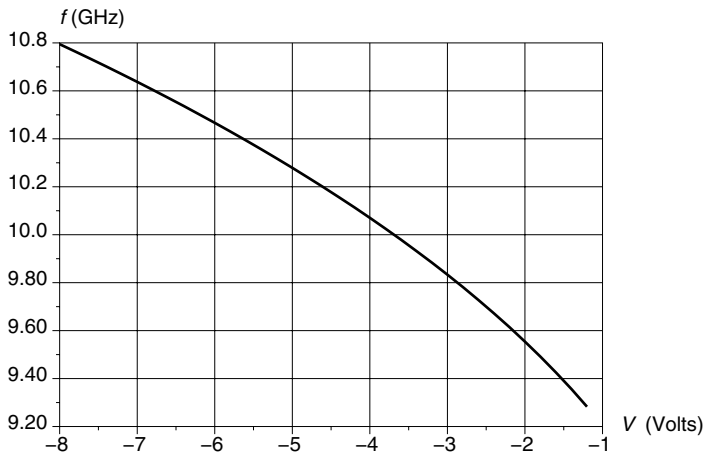


FIGURE 11.14

Frequency of the oscillations of the VCO versus the control voltage.



FIGURE 11.15

Amplitude of the oscillations of the VCO versus the control voltage.

reads:

$$v_c(t) = g(\eta(t - T)) = \beta \sin^2(\alpha \cdot (\eta(t - T) - \eta_0) - \varphi_0) + V_{c0} \quad (11.22)$$

where β, α, φ_0 , and V_{c0} are parameters that can be adjusted to produce the desired chaotic signal. The parameter β plays a special role as it controls the degree of feedback that is applied to the loop. For $\beta = 0$, the oscillator acts as a purely sinusoidal one and the frequency is adjusted through the value of V_{c0} . The loop filter has a time constant τ .

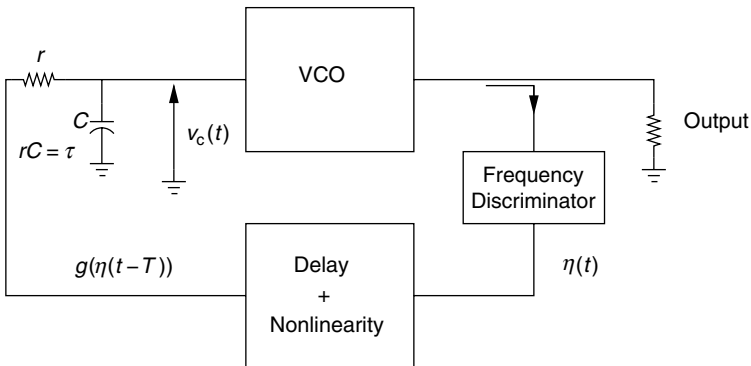


FIGURE 11.16

Chaotic oscillator $T/\tau = 64 V_{c0} = -4$ V.

Introducing the equation of the feedback loop in the general envelope transient equations of the oscillator leads to the following system:

$$\begin{aligned} \frac{d\bar{Q}(\bar{X}(t_e))}{dt_e} + \eta(t_e) \cdot \bar{\Omega}_0 \bar{Q}(\bar{X}(t_e)) + \bar{F}(\bar{X}(t_e)) + \bar{Y}_s \cdot (\bar{E}_s - \bar{X}(t_e)) &= 0 \\ \bar{Y}_s \cdot (\bar{E}_s - \bar{X}(t_e)) &= 0 \\ \tau \frac{dV_r(t_e)}{dt_e} + V_r(t_e) - g(\eta(t_e) - T) &= 0 \end{aligned} \quad (11.23)$$

System (11.22) has to be discretized using a full implicit scheme, thus giving the discretized system (11.4.2) at time $t_e = k \cdot \Delta t_e$:

$$\begin{aligned} \left(\frac{1}{\Delta t_e} + \eta^{(k)} \cdot \bar{\Omega}_0 \right) \bar{Q}^{(k)}(\bar{X}^{(k)}) + \bar{F}^{(k)}(\bar{X}^{(k)}) + \bar{Y}_s \cdot (\bar{E}_s^{(k)} - \bar{X}^{(k)}) \\ - \frac{\bar{Q}^{(k-1)}(\bar{X}^{(k-1)})}{\Delta t_e} = 0 \\ \bar{Y}_s \cdot (\bar{E}_s^{(k)} - \bar{X}^{(k)}) = 0 \\ \left(1 + \frac{\tau}{\Delta t_e} \right) V_r^{(k)} - g(\eta(k \cdot \Delta t_e) - T) - \frac{\tau}{\Delta t_e} V_r^{(k-1)} = 0 \end{aligned} \quad (11.24)$$

The discretized system has been solved for different values of the parameters using a software written with Scilab [21]. It gives the time evolution of the oscillation frequency and amplitude. It has to be noticed that the chaotic oscillator is an autonomous system relative to the high-frequency time scale as well as to the low-frequency time scale.

To investigate the full behavior of the chaotic oscillator, the bifurcation diagram can be calculated as the parameter β varies. This bifurcation diagram is plotted by sampling the oscillation frequency at instants multiple of the delay T introduced in the loop. Such diagrams are plotted in Figure 11.17 and Figure 11.18, for different values of the parameter α . Inspecting the plot given in Figure 11.17, for $\alpha = 314rd$ shows that for the values $\beta < 0.17V$ the oscillator works as a pure oscillator with a frequency decreasing as β increases. Above this value, the steady state regime bifurcates in a two frequency regime between the reduced frequencies

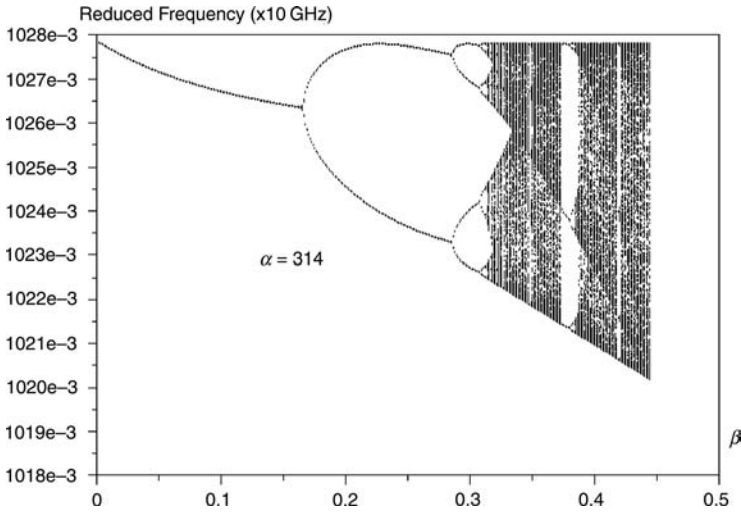


FIGURE 11.17
Bifurcation diagram of the chaotic oscillator for $\alpha = 314rd$.

$\eta_1 = 1.023$ and $\eta_2 = 1.028$. This regime is illustrated in Figure 11.19 where the transient and the steady state regimes are plotted. It can be noticed that starting from a quasi-chaotic transient the working regime of the oscillator stabilizes to a two-frequency FSK steady state regime (*d*).

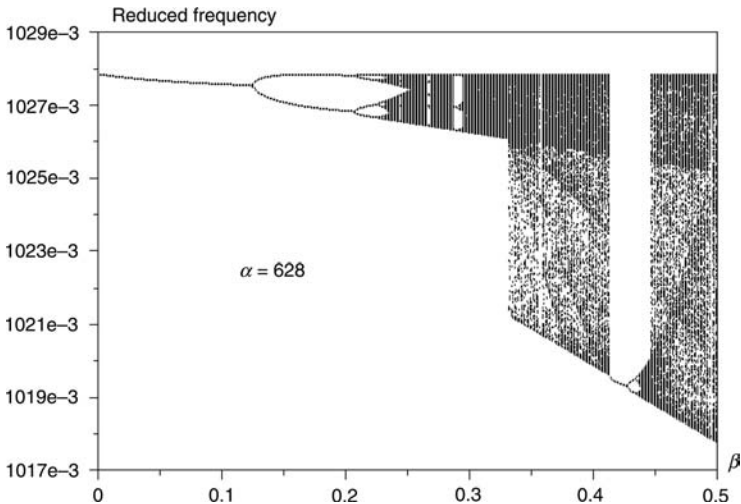


FIGURE 11.18
Bifurcation diagram of the chaotic oscillator for $\alpha = 628rd$.

Further increasing β leads to a second bifurcation which produces a four-frequencies regime and so on, up to the birth of a chaotic regime. For some values of β , regimes with three or five frequencies arise which are typical of a subharmonic route to chaos.

One advantage of the analyzed chaotic oscillator is the ability to control the bandwidth of the frequency spectrum by adjusting the value of the parameter β . Indeed if we examine the spectrum of the chaotic oscillator shown in Figure 11.20, one can notice that for the chosen value of β

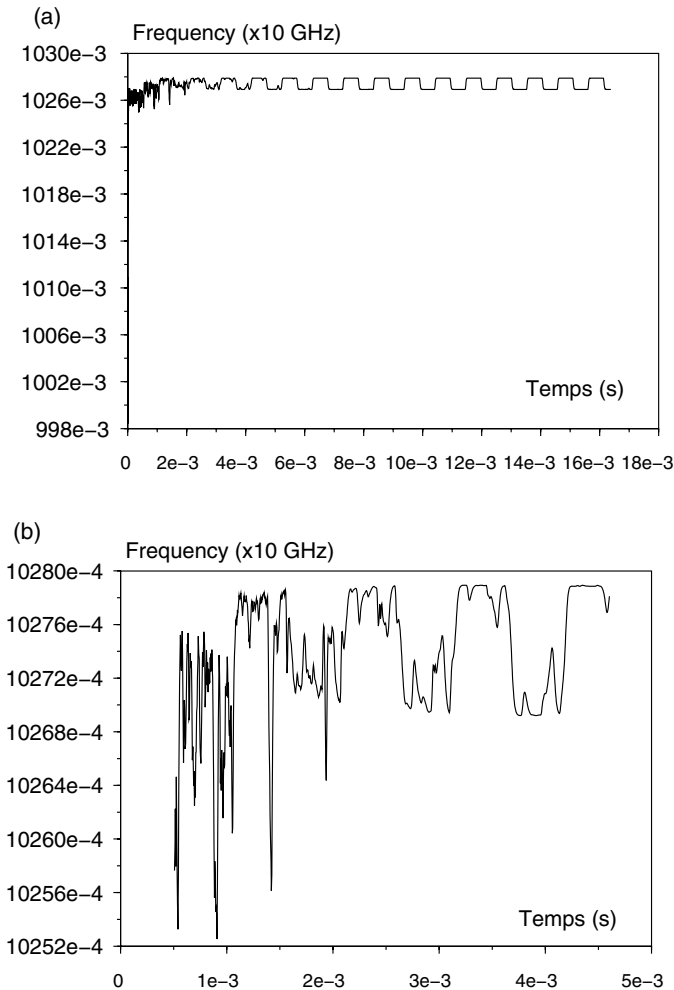
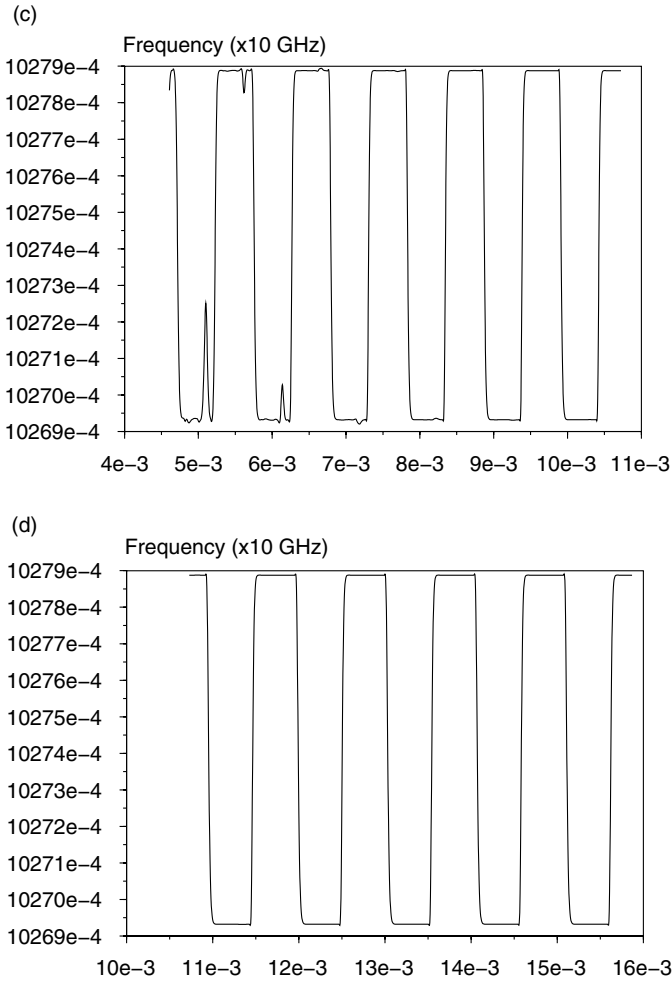


FIGURE 11.19
Transient set-up of a two-frequencies steady state regime.

**FIGURE 11.19***Continued.*

(i.e., $\beta = 0.36$) the spectrum bandwidth is given approximately by the peak to peak deviation of the reduced frequency times the normalization frequency.

$$BW \approx \Delta\eta \cdot f_0$$

This result corroborates the Carson rule for frequency modulated signals [22].

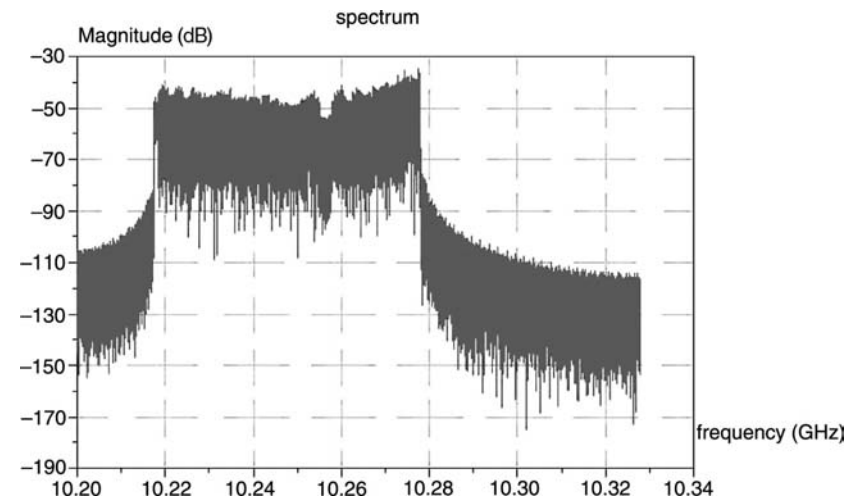


FIGURE 11.20
Spectrum of the output signal in the chaotic regime.

11.5 Chaotic Modulator and Demodulator

The previous chaotic oscillator has been used to transmit information signal $m(t)$ by adding it to the feedback signal generated by the nonlinear network as shown in Figure 11.21. In this case, the information signal participates in the dynamics of the oscillator. The control signal of the VCO now satisfies

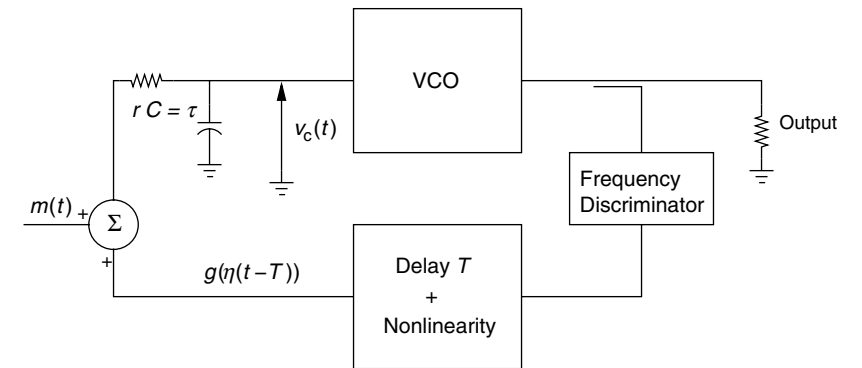


FIGURE 11.21
Architecture of the chaotic modulator.

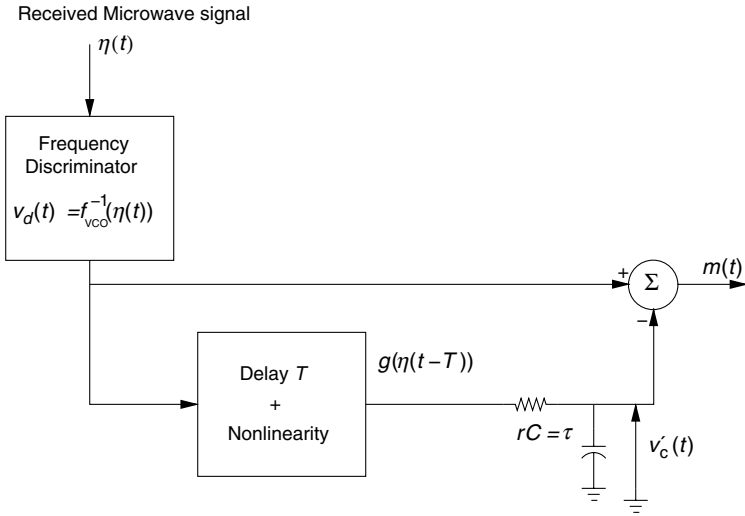


FIGURE 11.22
Structure of the chaotic receiver (demodulator).

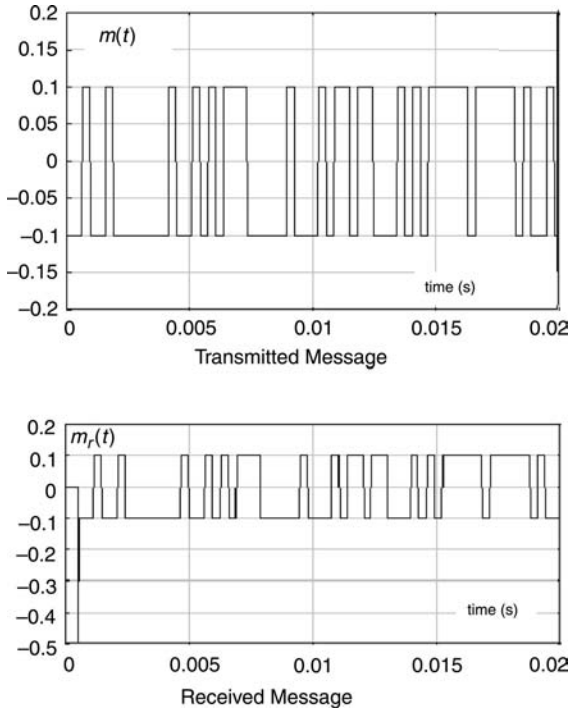


FIGURE 11.23
Transmitted and received signals in the case of a perfect match between parameters.

the equation:

$$\tau \frac{dv_c(t)}{dt} + v_c(t) = g(\eta(t - T)) + m(t) \quad (11.25)$$

where

$$\eta(t) = f_{VCO}(v_c(t)) \quad (11.26)$$

Under the same conditions as those described in Section 11.4, the frequency of the VCO varies in a chaotic way and the information signal is embedded in the dynamics of the process.

The structure of the receiver is shown in Figure 11.22. It consists of a frequency discriminator which has to produce a voltage $v_r(t)$, image of the reduced frequency of the received signal:

$$v_r(t) = f_{VCO}^{-1}(\eta(t)) \quad (11.27)$$

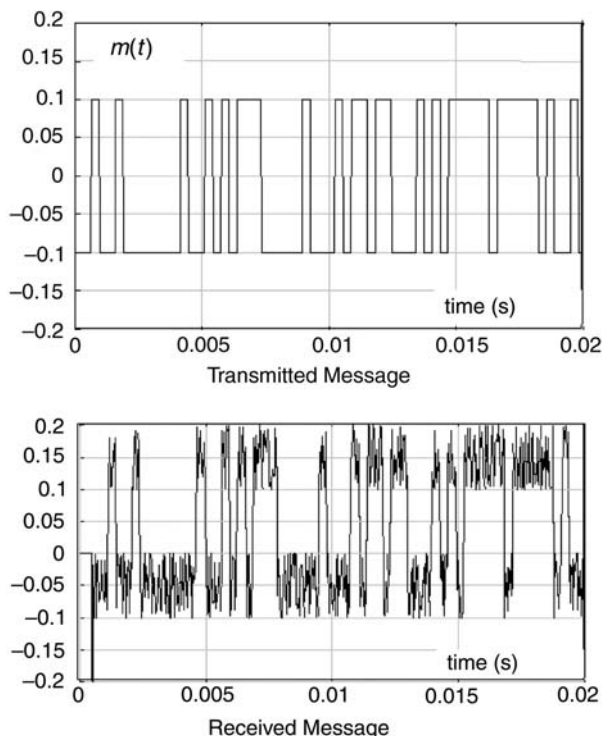


FIGURE 11.24

Transmitted and received signals in the case of a 10% mismatch between parameters.

This frequency discriminator can be produced with a Phase Locked Loop (PLL) using the same VCO as the one used in the emitter part of the MODEM. The voltage $v_r(t)$ is applied to the delayed nonlinearity giving:

$$\tau \frac{dv'_c(t)}{dt} + v'_c(t) = g(v_r(t - T)) \quad (11.28)$$

Thus subtracting the voltages $v_r(t)$ and $v'_r(t)$ gives the received information message $m(t)$. Synchronization time is equal to the delay time T .

Using the model of the chaotic VCO given in the previous section, the chaotic modulator–demodulator can be fully simulated to check the robustness of synchronization versus the parameters mismatches. In Figure 11.23, the comparison of emitted and received information signals are given in the case of a perfect match between parameters. A change of 10% in the receiver parameter β leads to the situation given in Figure 11.24 where the received signal becomes very noisy. Moreover, the BER of such a transmission can be calculated using the transient envelope simulation to give the curve shown in Figure 11.25, where it can be noticed that the BER can be maintained in a range of about 12% for parameter α .

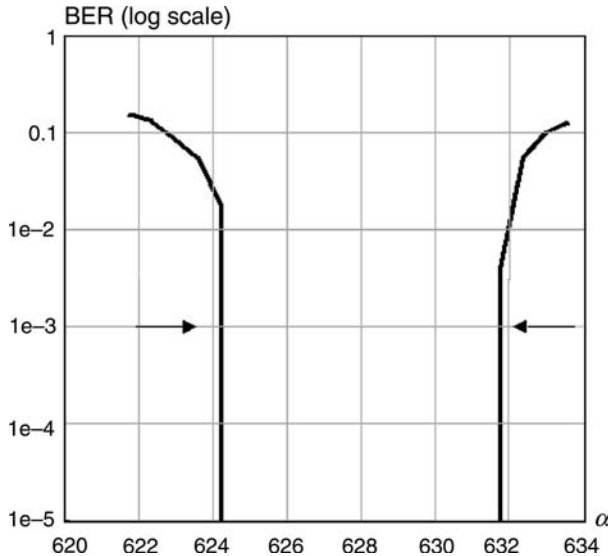


FIGURE 11.25

The BER of the chaotic modulator–demodulator.

References

1. L.M. Peccora and T.L. Carroll, Driving systems with chaotic signals, *Phys. Rev. A*, 44 (4), 2374–2383, 1991.
2. T.S. Parker and L.O. Chua, *Practical Numerical algorithms for Chaotic Systems*, Springer-Verlag, Berlin, 1989.
3. G. Kolumban, M.P. Kennedy and L.O. Chua, The role of synchronization in digital communications using chaos—part II: chaotic modulation and chaotic synchronization, *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, 45 (11), 1129–1140, 1998.
4. H. Nijmeijer and I. Mareels, An observer looks at synchronization, *IEEE Trans. Circuits Syst. I*, 44 (10), 882–890, 1997.
5. Y. Liu and P. Davis, Dual synchronization of chaos, *Phys. Rev. E*, 61 (3), 2176–2179, 2000.
6. M. Hasler, Synchronization of chaotic systems and transmission of information, *Int. J. Bifurcat. Chaos*, 8 (4), 647–659, 1998.
7. F. Bohme and W. Schwarz, The Chaotizer–Dechaotizer channel, *IEEE Trans. Circuits Syst. I*, 43 (7), 596–599, 1996.
8. G. Mazzini, G. Setti and R. Rovatti, Chaotic complex spreading sequences for asynchronous DS-CDMA—Part I: system modeling and results, *IEEE Trans. Circuits Syst. I*, 44 (10), 937–947, 1997.
9. Heidari-G. Bateni and C.D. McGillem, A chaotic direct-sequence spread spectrum communication system, *IEEE Trans. Commun.*, 42 (2–4), 1524–1527, 1994.
10. T. Yang and L.O. Chua, Chaotic digital code division multiple access communications systems, *Int. J. Bifurcat. Chaos*, 7 (12), 2789–2805, 1997.
11. F. Dachsel and W. Schwarz, Chaos and cryptography, *IEEE Trans. Circuits Syst. I*, 48 (12), 1498–1509, 2001.
12. S. Penaud and P.H. Bouysse, DSP implementation of self-synchronised chaotic encoder decoder, *Electron. Lett.*, 36 (4), 365–366, 2000.
13. L. Larger, J.P. Goedgebuer and F. Delorme, Optical encryption system using hyperchaos generated by an optoelectronic wavelength oscillator, *Phys. Rev. E*, 57 (6), 6618–6624, 1998.
14. G. Kolumban and M.P. Kennedy, The role of synchronization in digital communications using chaos—part III: performance bounds for correlation receivers, *IEEE Trans. Circuits Syst. I*, 47 (12), 1673–1683, 2000.
15. C.P. Silva and A.M. Young, Implementations and uses of high frequency chaotic oscillators, in *Proceedings of the WS7 Nonlinear Phenomena in Microwave Electronic Circuits and Chaos: Analysis and Applications*, 30th EuMC, Paris, October, 2000.
16. E. Ngoya and R. Larcheveque, Envelope transient analysis: a new method for the transient and steady state analysis of microwave communication circuits and systems, in *Proceedings of the IEEE MTT'96 TH2B2*, 1996.
17. R. Quéré, E. Ngoya, M. Camiade, A. Suárez, M. Hessane and J. Obregon, Large signal design of broadband monolithic microwave frequency dividers and phase-locked oscillators, *IEEE Trans. Microwave Theory Tech.*, MTT-41 (11), 1928–1938, 1993.

18. H.G. Brachtendorf, G. Welsch, R. Laur and Bunse-A. Gerstner, Numerical steady state analysis of electronic circuits driven by multi-tone signals, in *Electrical Engineering*, Vol. 79, Springer Verlag, 1996, pp. 103–112.
19. J. Guittard, S. Penaud, J.C. Nallatamby and R. Quéré, Full analysis of a chaotic microwave oscillator for use in a FM-CSK communication system, *Int. J. RF Microwave Comp. Aided Eng.*, 12 (5), 469–474, 2002.
20. J. Guittard, Techniques d'analyse d'oscillateurs chaotiques: application aux télécommunications par synchronisation de chaos, PhD Thesis, Université de Limoges, France.
21. Scilab available at <http://www-rocq.inria.fr/scilab/>.
22. R.E. Ziemer and W.H. Tranter, *Principle of Communications*, Houghton Mifflin Company, Boston, 1976.

12

Chaos, Optical Systems, and Application to Cryptography

L. Larger

CONTENTS

12.1 From Chaos to Cryptography through CDMA	454
12.1.1 Basic CDMA Principles	454
12.1.2 Similarities with Chaotic Behavior	455
12.1.3 A Typical Transmission System Using Chaos Encryption	456
12.2 Which Dynamical System for Encryption?	457
12.2.1 Required Properties	457
12.2.2 Nonlinear Dynamics Classification	457
12.2.3 Nonlinear Delayed Differential Dynamics	458
12.3 NLDDE Performed in Optics and Optoelectronics	461
12.3.1 The Ikeda Ring Cavity	461
12.3.2 Physical Principles Involved in the NLDDE Set-Up	462
12.3.2.1 Amplitudes	462
12.3.2.2 Times	463
12.3.3 Optoelectronic Set-Ups Using the Laser Wavelength	463
12.4 Coding–Decoding Information in Chaos	467
12.4.1 Masking Techniques	467
12.4.2 Synchronization and Decoding Principles	468
12.4.3 Experimental Results	470
12.5 Confidentiality and Cryptanalysis Approach	472
12.5.1 Key-Space Dimension	472
12.5.2 Dynamics Identification from Time Series	472
12.6 Other Optical Set-Ups for Encryption	473
12.6.1 Chaotic Laser Intensity Using Electro-Optic Devices	473
12.6.2 Chaotically Coherence Modulated Light Beam	474
12.6.3 “All-Optical” External Cavity Laser Diode	475
12.6.4 Direct Optoelectronic Feedback in SC Lasers	476

12.7 Conclusions	477
References	478

12.1 From Chaos to Cryptography through CDMA

Chaos might be considered, from different points of view, as a fascinating subject of research, as a perturbing phenomenon that one wants to get rid off, or also as an interesting deterministic behavior whose properties can be of great interest if it can be understood and controlled. Regarding the last situation, chaos-based encryption has become one of the most important fields of applied research of the chaos theory in the last decade, since it appeared to be a promising and original practical application, as well as a sought after and in-demand field of research in the current communication- and information-driven society. More and more transmitted data are indeed requiring a certain level of confidentiality: private mobile phones, Internet transactions, medical data transfer, bank information exchange, etc. Cryptography has moved very recently from being quite a limited field of concern (diplomacy, national security, and military applications) to a very wide area concerning nearly everyone, thus requiring a high bandwidth encryption capability.

Although *code division multiple access* (CDMA) and chaos-based encryption are communication techniques that have been developed independently at very different times, one could easily have thought they were tightly connected due to the similarity existing in the involved principles. This similarity is actually so strong that a current research topics in chaos theory applications concerns the improvement of CDMA efficiency through the use of chaotic dynamic (see Chapter 11). In this section, we will first describe briefly the principles involved in CDMA, and then we will emphasize on the similarities between CDMA and encryption using chaos, as an introduction to chaos-based encryption principles.

12.1.1 Basic CDMA Principles

CDMA is a spread spectrum digital communication technique, meaning practically that the spectrum of the transmitted binary signal (carrier with the information) is much wider than the one of the information. At first sight, this should not represent a very efficient character of this particular transmission technique. However, it is also possible to transmit simultaneously and on the same bandwidth, different channels, thus reducing the

effective frequency cost per channel. This particular property of possible time–frequency coexisting channels is related to the fact that the carrier, in contrast to conventional modulation techniques based on sine waveforms, consists in a pseudo-random binary sequence (of finite length): it forms the code defined for one channel. Each bit of the message is XORed with the pseudo-random bit sequence, thus performing the CDMA modulation: the random bits being much faster than the message bits, the message spectrum is at the same spread through the code randomness, as well as shifted toward higher frequencies corresponding to the code bit frequency.

The demodulation process involves a second logical XOR operation, with the same pseudo-random binary sequence, the code. At this stage, the code appears to be equivalent to a key: the demodulation is indeed quite a hard task if the code used at the modulation process is not known for demodulation. In other words, the binary sequence provides a carrier means as well as an *encryption* means at the same time.

Complex theoretical criteria allow the definition of an optimal deterministic codes generation method, corresponding to the so-called “Gold-sequence.” This sequence is a huge random binary sequence, from which each code of each channel is picked. Typical applications of CDMA are GPS signals and the modulation techniques of the third-generation mobile phones.

12.1.2 Similarities with Chaotic Behavior

Chaotic dynamics should now appear more obviously as potential candidates for pseudo-random carriers, providing both an information carrier capability for transmission as well as an encryption function. Through their low predictability or sensitivity to initial conditions they exhibit the random character of interest. In other words, they exhibit wider spectra when compared with traditional periodic waveforms, thus fulfilling the spread spectrum nature earlier described for CDMA.

In contrast, chaotic dynamics are intrinsically deterministic since they are generated by differential equations, similarly to the fact that the CDMA codes are practically produced by well-defined and optimized algorithms (e.g., the Gold sequence). The determinism is of great importance when designing the decoder, involving, for example, synchronization techniques between local and distant chaotic waveforms. CDMA decoding techniques also involve such synchronization procedures.

However, the analogy between CDMA and the chaotic optical encryption systems of concern here, should not be taken too far, as there are currently still some fundamental differences between them. CDMA is an older and quite mature communication technique and it is already applied on a wide scale. CDMA is mainly a digital technique, whereas chaos-based

encryption system deals mainly with analog signals. The chaotic sequences of concern are of nearly infinite length when compared with the finite length of CDMA codes.

Chaotic secure communication systems do not yet pretend to compete with CDMA, as one will notice that the systems reported in the next sections are still under development. We will now concentrate on chaos-based encryption in optics.

12.1.3 A Typical Transmission System Using Chaos Encryption

Figure 12.1 represents the bloc diagram of a typical information transmission system secured by chaotic waveforms. The emitter used by Alice consists of a chaos generator providing the chaotic carrier and a mixing bloc whose role is to hide the information message inside the chaotic waveform. The resulting signal is transmitted to the authorized receiver, Bob, who owns an adequate decoder. That decoder consists of a second nearly identical chaos generator and a synchronization bloc. Synchronization is the most critical step in the decoding procedure. If the locally generated chaos is synchronized with the one received from the emitter, the extraction of the information is then usually performed through quite a straightforward operation, typically a subtraction. Of course, since we are claiming the transmission is secure, the figure also takes into account that the line might be tampered by an unauthorized person, Eve (the eavesdropper). The following sections will detail nearly in the same order the different operations just described. We will first discuss the chaos generator chosen, and its particular dynamical properties in the context of encryption.

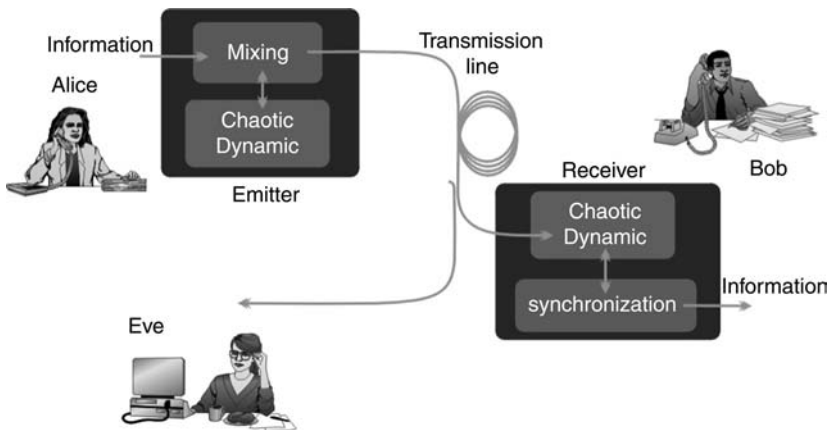


FIGURE 12.1

Typical transmission system using chaos encryption.

Practical realizations of the chaos generator in optoelectronics will be reported, as well as experimental characterization. Coding and decoding problems will then be addressed by illustrations with experimental results. Finally some notions on security and cryptanalysis will be discussed before concluding and reporting some current research on other optical systems for chaos-based encryption systems.

12.2 Which Dynamical System for Encryption?

First of all, we will describe and justify the particular type of dynamical systems that have been chosen in most of the chaos-based optical encryption systems.

12.2.1 Required Properties

Among the numerous experimental set-ups capable of generating chaos, one has to choose the most efficient one for encryption purposes. Of course, the chaotic dynamic should be as complex as possible, so that, for example, it can be very difficult for Eve to analyze it and find out the underlying determinism. In contrast, the set-up should not be too complicated, so that designing an efficient decoder is not a task too hard, assuming the “key” parameters are known. A simple design is also sometimes required for stability properties of the encoding and decoding process.

12.2.2 Nonlinear Dynamics Classification

Table 12.1 is a kind of classification [1] of various dynamical systems, according to their discrete, continuous, invertible, and noninvertible nature (rows), and also with respect to their phase space dimension (columns). Each box of the table is filled with the most complex possible dynamical behavior corresponding to a given row and a given column.

TABLE 12.1
Dynamics Classification

	1	2	3	...	∞
Flow	Fixed Point	Torus	Chaos	...	Chaos
Invertible Map	Torus	Chaos	...	...	Chaos
Noninvertible Map	Chaos	...	...	...	Chaos

According to this table, the criteria of a complex behavior would correspond to last column (high-dimensional phase space), and the simplest structure that would still be able to exhibit a chaotic behavior would be represented by the first column and last row (noninvertible map in one-dimension). Although both criteria seem to be difficult to fulfill simultaneously, there exists a particular and unusual class of dynamical systems which might exhibit features in terms of behavior complexity as well as system simplicity: the nonlinear delay differential dynamics. This kind of dynamics is located precisely in the last column in Table 12.1 due to the presence of a delay. Surprisingly, however, the minimum number of physical dynamical variables required to produce such a chaos can be as low as 1, as it is for a discrete mapping. It is also important to notice that the dynamics is a flow, thus corresponding to a real-world physical system and not a numerical one.

12.2.3 Nonlinear Delayed Differential Dynamics

A typical mathematical writing of that particular dynamics is given in Equation (12.1):

$$\frac{dy}{dt}(t) = f[y(t), y(t - T_1), y(t - T_2), \dots], \quad (12.1)$$

where y is the physical dynamical variable (it can be scalar or vectorial), and $f[\cdot]$ is a nonlinear transformation depending on y at time t , but also on y at the moments delayed in time by the quantities T_i ($i = 1$ to n , and T_i are ordered with decreasing values of the delay). Regarding state space dimension, it is now clear that it has to be infinite according to the initial conditions required to uniquely determine a phase space trajectory: one would require not only the components (in finite number) of the initial value $y(t_0)$ but also all (in infinite number) the continuous time variations of those components on a time length T_1 (the largest delay).

When considering only a scalar dynamical variable (first-order dynamics) and a single time delay involved in the arguments of the nonlinear function, we obtain the simplest form of a *nonlinear delay differential equation* (NLDDDE):

$$\tau \frac{dy}{dt}(t) + y(t) = \beta \cdot f[y(t - T)] \quad (12.2)$$

where τ is a time scaling factor representative of a physical response time of the real system. In the following, we will only consider the large delay case, for which T is large when compared with τ . Although the small or comparable delay case might also be of interest in other situations, in our encryption application we are seeking a dynamical system capable of

high-complexity chaotic behavior whose phase space dimension is related to the delay (size of the initial conditions). It is important here to notice that in the large delay case, one can obtain, with such NLDDs, very high complexity chaotic behaviors, with attractor dimension as high as several hundreds (to be compared with the dimensions lower than 3 for the attractors of the Lorenz type).

In Equation (12.2), the delayed feedback nature of the dynamics appears more clearly. The linear differential operator in the time domain $[\tau d/dt + Id]$ is indeed the consequence of a first-order filter with the well-known Laplace transform $1/(1 + \tau p)$. The bloc diagram of the physical system modeled by Equation (12.2) is depicted in Figure 12.2.

It consists of a pure delay line generating a time delay T constant for each frequency (no dispersion), a nonlinear transformation $f[\cdot]$, a linear low pass filter with a cut-off frequency $f_c = 1/2\pi\tau$, and a linear amplifier with a gain β . The gain β is a multiplicative factor for the nonlinear function, thus acting in the feedback as a weight of the nonlinear transformation. In many situations, β is used as a bifurcation parameter, since it measures the importance of the nonlinear action in the feedback process. For low values of β , the system can be linearized, thus recalling a basic harmonic oscillator, for which the gain condition can be fulfilled through β , and the phase condition depends on the delay. Because of the large delay, the phase condition is easily met for many frequencies, and hence, the most crucial point to start an oscillation from a stable steady state is to have enough gain. While increasing β , the nonlinear transformation is more and more important in the actual dynamic, till the chaotic regimes. As depicted in Figure 12.3, a typical period doubling route to chaos is observed in the bifurcation diagram, while increasing β from zero (the horizontal axis is β , the vertical one corresponds to the visited amplitudes of y for a given dynamical regime, the gray scale encodes the probability density of y).

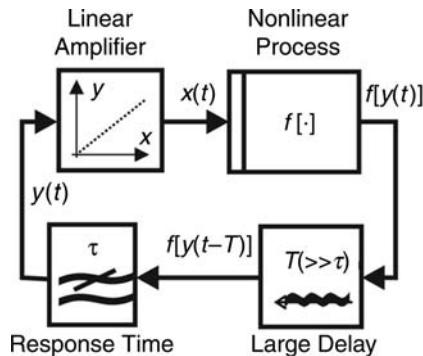
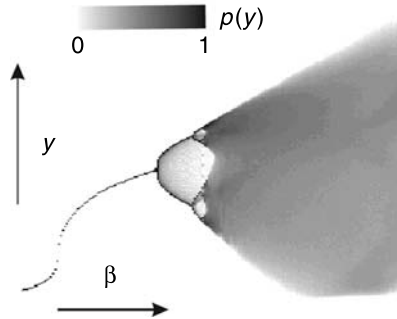


FIGURE 12.2

Bloc diagram of the scalar nonlinear delayed dynamic.

**FIGURE 12.3**

Bifurcation diagram calculated from Equation (12.2).

Another qualitative approach sometimes used for the scalar NLDDE involves the so-called adiabatic approximation. Although this approach cannot be quantitative, it might be of interest while looking for the first time at the instability phenomenon that might occur for a given nonlinear function $f[\cdot]$. If we assume in Equation (12.2) that time derivatives can be neglected most of the time, then the system reduces to a mapping, since we have $y(t) = \beta \cdot f[y(t - T)]$. Supposing the trajectory corresponds to constant values of y on successive time intervals of duration T , and denoting this value on the n th interval y_n , we obtain the mapping $y_n = \beta \cdot f(y_{n-1})$.

When $\beta f[\cdot]$ is replaced by the reverse parabola, a mapping corresponding to the logistic application is obtained. Exploring the different steady states and the values reached during any periodic behavior are then found classically, when studying the fixed points of the p -iterated function $\beta f[\cdot]$. The main difference from the logistic map is that in the practical cases corresponding to the experiments, the function $f[\cdot]$ exhibits several extrema. This is of great importance in terms of the complexity of the actually observed dynamical regimes.

The adiabatic approximation (Figure 12.4) is practically valid only for low complexity regimes, like the steady states, the $2T$ and $4T$ periodic regimes, for which the experimental behavior corresponds effectively to square waveforms between plateaus (on which the condition $[d/dt \text{ small}]$ is easily verified). For higher values of β (i.e., for higher complexity regimes), the transients where the time derivative reaches large values plays an important role, and the regimes obtained with the mapping are significantly different from the ones obtained with the flow (NLDDE). As an example, the period-3 window observed on the mapping in Figure 12.5 is not present (at least at the same position) in the case of the continuous time NLDDE bifurcation diagram in Figure 12.3.

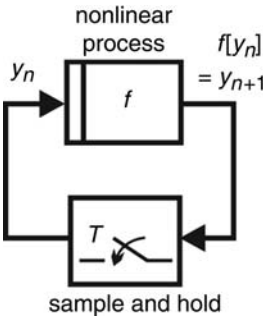


FIGURE 12.4
Bloc diagram in the adiabatic approximation situation.

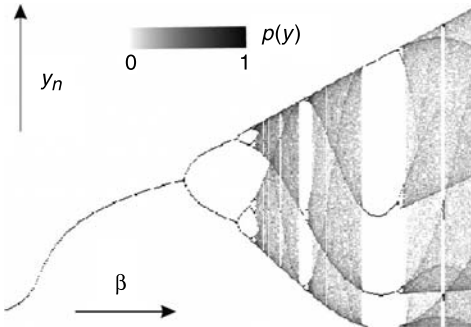


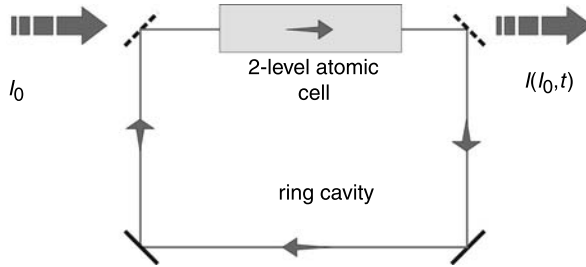
FIGURE 12.5
Bifurcation diagram calculated from a mapping using $\beta f[\cdot]$.

12.3 NLDDE Performed in Optics and Optoelectronics

Following the scheme described in the previous section, we developed different experimental set-ups intended to perform chaos generators described by NLDDEs. All of these set-ups are inspired by a brain experiment initially proposed and numerically explored by Kensuke Ikeda in the early 1980s [2].

12.3.1 The Ikeda Ring Cavity

The experiment is depicted in Figure 12.6. The feedback is performed all optically in a ring cavity with two partially reflecting mirrors (input and output mirrors in the top of the figure) and two totally reflecting mirrors (the two mirrors in bottom of the figure). The nonlinear transformation

**FIGURE 12.6**

The Ikeda ring cavity.

can be performed differently through the interaction of the light beam inside the cavity with a two-level atomic cell. One way is to benefit from a refractive index change in the cell with respect to the light intensity (thus changing the total optical path viewed by the beam inside the cavity: the so-called Kerr effect). The light intensity is actually determined by the interference between the input light beam and the feedback light beam. But the later interference condition depends also, through the feedback, on the total optical path in the cavity determined by the previous refractive index change in the cavity. The nonlinear function then corresponds to the interference figure between two waves having a phase difference Φ ; the interference intensity is thus a cosine function with respect to the phase difference Φ generated by the optical path through the cavity. The iteration process occurs with a time scale corresponding to the traveling time of the beam through the cavity; the delay is thus generated via a traveling wave through a medium of a fixed length.

The observed quantity is the output intensity resulting from the interference in the cavity at time t . Ikeda showed that this output intensity has a dynamic which can be described, under some assumptions, by an equation of the same type as Equation (12.2).

An important parameter is the constant input laser intensity I_0 , which determines how strong the light-matter interaction in the atomic cell can be. For low input intensity, a stable steady state is observed, while period doubling appears for an increasing input intensity. For high enough intensity, the output becomes chaotic.

12.3.2 Physical Principles Involved in the NLDDE Set-Up

12.3.2.1 Amplitudes

Let us go back to the bloc diagram in Figure 12.2. Instead of considering the linear element between x and y as an amplifier, one could treat it as a conversion bloc between two physical variables of different nature. The

systems we are considering in optics are not only described by an optical intensity or a refractive index but also wavelength, optical phase, or optical path difference might be involved in the physical set-up; even voltage or current is of concern with optoelectronic systems. This difference in the nature of the variables is actually imposed by the physics involved in the set-up: we want a highly nonlinear behavior as well as a feedback system which reacts on itself after a delay (the latter implies that the physical nature of the feedback signal has to be the same as the input of the nonlinear function). More precisely, the linear conversion bloc from x to y is performed by the Kerr effect which relates the intensity (x) and the refractive index change, or equivalently the optical path difference change (y); the nonlinear transformation occurs precisely between the same variables, since the cosine output is the intensity (x) whereas the input is the optical path change (y).

12.3.2.2 Times

As already noticed in the equations, the time scales are dual. One is related to the delay generated by the traveling wave (approx. 1 or 2 m in the air, hence $T \simeq 3 - 10$ nsec), and the other one corresponds to the slowest physical response time involved in the feedback. In the optical ring cavity, this limitation occurs on very short time scales τ , since light-matter interaction in the atomic cell can be as fast as a few picoseconds or femtoseconds. The Ikeda cavity then corresponds effectively to the large delay case of the NLDDE systems. In the latter case, the presence of very short time scales is one of the main reasons why the Ikeda set-up is mainly a mental experiment, since signals would have been too fast to be properly managed; another is that the optical intensity required for a significant Kerr effect has to be actually very high compared to conventional optical power levels.

12.3.3 Optoelectronic Set-Ups Using the Laser Wavelength

In the mid-1990s, our group explored the possibility of experimentally performing the Ikeda ring cavity, but involving unusual optoelectronic means: with the recently available wavelength tunable semiconductor laser. The initial idea [3] was to modulate an optical interference situation not through the refractive index change (as in the Ikeda experiment) but through the wavelength change. The phase shift between two waves can be indeed written as $2\pi\Delta/\lambda$, where Δ is the optical path difference, and λ is the wavelength usually considered as a constant parameter. In our set-up, the aim is to obtain chaotic fluctuations of the wavelength of a laser $\lambda(t)$. The corresponding set-up is depicted in Figure 12.7 [4].

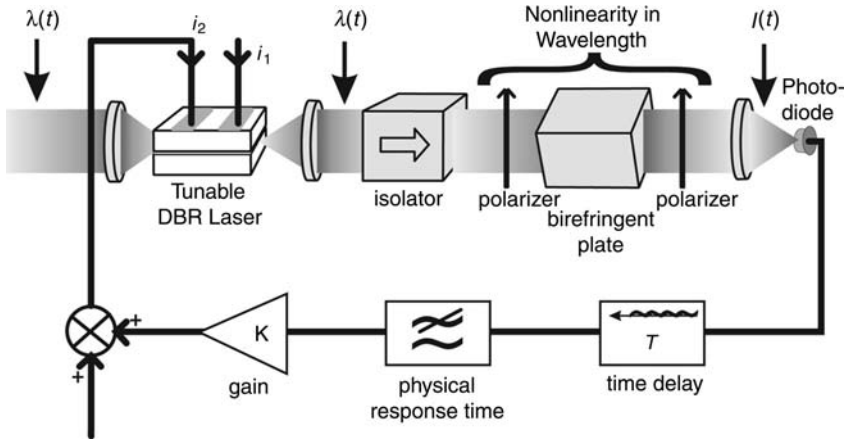


FIGURE 12.7

The wavelength chaos generator.

The main element of the wavelength chaos generator is the tunable two-section DBR semiconductor laser. One electrode adjusts the laser power through the current i_1 as usual, and the other can be used to tune continuously and linearly the wavelength around $1.55 \mu\text{m}$ through another current i_2 , over a 2 nm range. This particular source converts an input current i_2 (x) linearly into an output wavelength deviation $\delta\lambda$ (y) around the central wavelength λ_0 . Next, the nonlinear transformation is performed similarly to the Ikeda ring cavity case, with a two-waves interferometer, performed here by a birefringent plate placed between two crossed polarizers. This interferometer nonlinearly converts the input wavelength fluctuation into an optical intensity fluctuation, according to a sine law, typically:

$$f[y] = \sin^2[y + \Phi_0] \quad (12.3)$$

where y is proportional to $\delta\lambda$ and Φ_0 is related to the central wavelength λ_0 . It is to be noticed that the system is not restricted, in principle, to the nonlinear function profile given in Equation (12.3), since any optical filter operating in the same spectral range and exhibiting at least one extremum can produce a chaotic behavior; the actual number of the sine extremum (Equation (12.3)) involved experimentally can be as high as 14, whose feature is important to obtain high-complexity chaotic dynamics. To perform the feedback, the output signal has to be compatible with the tuning current i_2 : this is fulfilled simply through the detection of the optical power fluctuations with a photodiode, thus producing a photo-current proportional to the nonlinear transformation.

The previous optoelectronic part of the oscillator allowed us to define most of the amplitude elements (linear conversion, nonlinear transformation, compatibility of the signals nature). The temporal

characteristics are determined by the electronic feedback. The first-order dynamical process, as written in the left-hand side in Equation (12.2), is performed by an active first-order low pass filter working in the audio range ($\tau \simeq 8.7 \mu\text{s}$). To perform a large delay NLDDE, we had to carry out a delay line with a long delay, typically $512 \mu\text{sec}$, thus producing a ratio T/τ of about 60. This was achieved electronically instead of using a very long fiber line; a CCD-based analog FIFO memory of 1024 elements was used for this, thus generating a delay that can be accurately tuned through an external clock frequency. A tunable electronic amplifier is also used in the electronic feedback as a means for adjusting the bifurcation parameter value (i.e., the overall weight of the nonlinear signal fed back to the tunable laser).

Numerous different dynamical regimes can be obtained while increasing (from the top to the bottom) the electronic feedback loop gain; some of them are reported with their FFT spectra in Figure 12.8. They correspond to different fixed values of the bifurcation parameter (the overall feedback gain). Periodic regimes are obtained for low values of β , with their typical line spectra. The first periodic regime observed while increasing β from zero is usually a square wave with a period of nearly $2T$, and the next stable periodic regime in the period doubling cascade gives rise to subharmonic

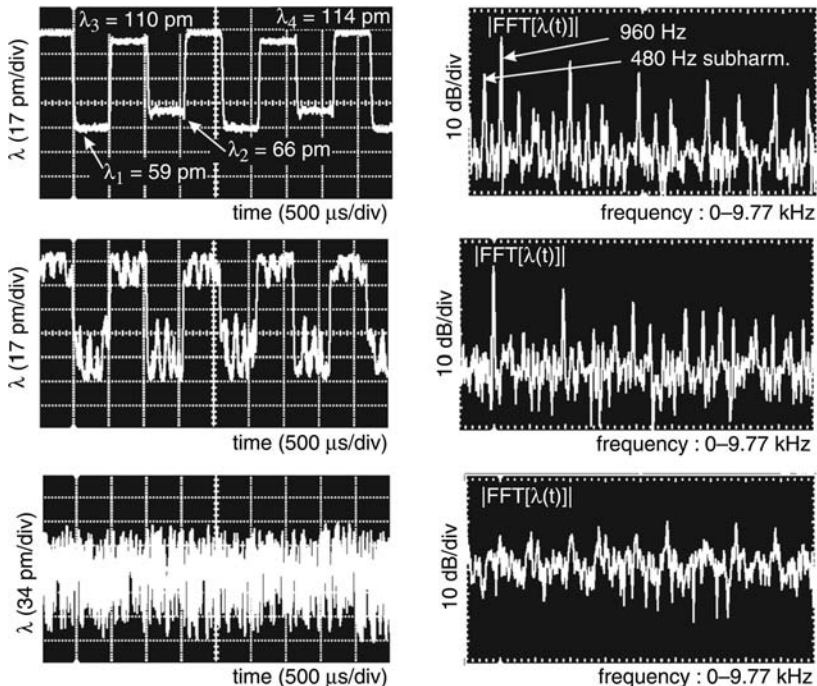


FIGURE 12.8

Experimental trajectories in time and frequency domain.

frequencies (period $4T$) as depicted in the first top experimental record; the next $2^n T$ periodic oscillations are difficult to observe experimentally due to noise and due to a small separation of the oscillation levels. After the period doubling cascade, a kind of inverse cascade can be recognized, consisting of a $2^n T$ carrier (n is now decreasing with increasing β) with small chaotic fluctuations on each level of the square carrier waveform. The middle trace in Figure 12.8 represents such a $2T$ carrier with the characteristic chaotic fluctuations superimposed to the plateaus. When β is large enough, the so-called fully developed chaos is obtained (bottom traces in Figure 12.8), where no periodic background can be seen in the spectrum, but only a high level noisy, and nearly flat, spectrum is observed within the entire oscillator bandwidth. Such chaotic regimes are of course of great interest in our application to encryption. They are obtained for many different values of the bifurcation parameter, provided it is sufficiently high.

When β is slowly increased in time, and when using an analog oscilloscope, it is possible to obtain an experimental bifurcation diagram in the same way as the ones obtained numerically in Figure 12.3 and Figure 12.5. The mean spot trace intensity along the vertical axis of the oscilloscope screen is then representative of the dynamical variable probability density. The resulting photography of such an oscilloscope screen is represented in Figure 12.9, for different values of the parameter λ_0 , the central wavelength

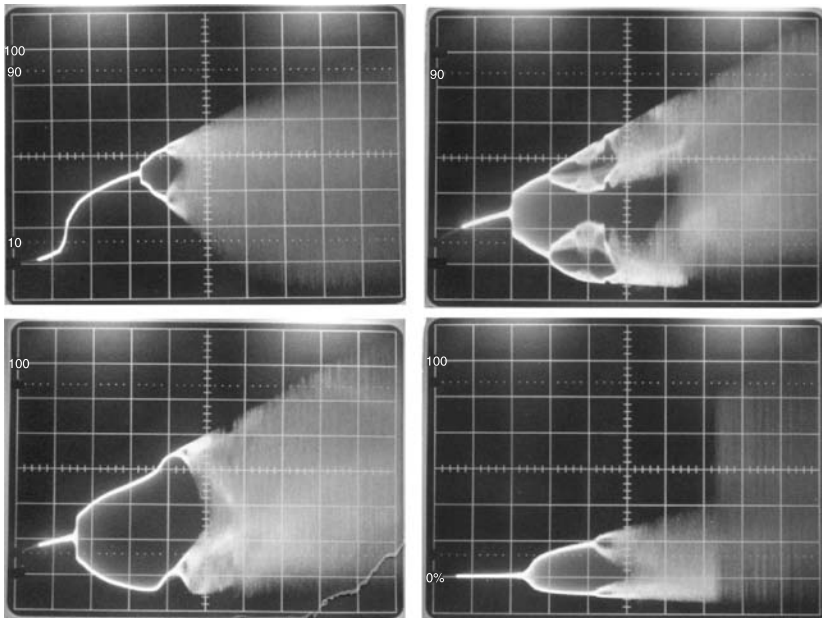


FIGURE 12.9

Experimental bifurcation diagrams.

of the laser. The first bifurcation diagram in the figure should be compared with the numerical one in Figure 12.3, since the experimental parameters were carefully adjusted to the same values as in the simulations.

Figure 12.9 shows that many different bifurcations can arise by slightly changing some parameters of the system. In the same vein, it is possible to generate many different chaotic regimes when choosing different parameter values, each of these constituting the key of the encryption process.

12.4 Coding–Decoding Information in Chaos

12.4.1 Masking Techniques

Most of the systems developed so far perform an encryption through an additive masking of the message inside the chaotic carrier. “Masking” means that the amplitude of the message is smaller than the one of the chaotic signal. The transmitted signal is thus nearly identical to the chaotic signal, which is chosen to appear as close as possible to a real noise; it should then be quite difficult for Eve to recognize any message in the transmitted signal. The first idea of the pioneering work of Pecora and Carroll [5] was to perform additive masking outside the chaotic oscillator (i.e., that the chaotic oscillation is generated completely independently of the message). The main drawback of that solution was that intrinsically the message had the same role as any additive noise introduced in the transmission channel. The technique is thus extremely sensitive to noise induced by the channel.

Another possibility is the one chosen in the wavelength chaos emitter in Figure 12.7, where the signal is added inside the oscillation loop (see the adder before the feedback on the laser tuning electrode). Since the message amplitude is small, it acts only as a small perturbation on the chaotic oscillation, which keeps its general noise-like properties in the presence of the message. The message insertion using an adder in the feedback loop is equivalent to the modulation of the parameter Φ_0 in Equation (12.3). Note that in principle any other parameter (e.g., β or T) of the dynamics can be chosen to mix the message with the chaotic carrier, but the decoding technique then strongly depends on it. For this particular reason, the technique for mixing the message with the chaos is sometimes called “chaos modulation.” In some situations, the transmitted message is binary, because there does not exist any demodulation technique to retrieve an analog message. The term CSK for chaos shift keying is then found in the literature [6]. We will next explain the decoding technique relative to the additive chaos modulation technique, which is the most common and the simplest one

for the demodulation, and for which both analog and digital decoding are possible.

12.4.2 Synchronization and Decoding Principles

The term “synchronization” is employed here because it is currently used in the literature, but it does not seem to be the correct term. A synchronization somehow means that a receiver owns at least a certain level of independence with respect to the emitter. In the situation described here, chaos replication [7] would be more appropriate, since the receiver is completely coupled to the signal transmitted by the emitter.

Assume first that no message is involved in the emitter. The proper decoder would consist of a nontrivial system giving at the output a null signal (no message), while using the chaotic waveform as the input, and generating from this signal, another one which is identical (i.e., synchronized or more rigorously replicated). The emitter–receiver replication principle is illustrated in Figure 12.10 using blocs similar to the ones used in Figure 12.2 (except for the linear filter, which is replaced by its impulse response, thus generalizing the low pass filter situation to any kind of linear filter). First it can be seen that the same elements are present at both sides. Second, the main difference relies on the fact that the receiver has an open-loop architecture and hence cannot produce any oscillation without external signal. The receiver could hence be called a passive nonlinear delayed dynamical process, whereas the emitter would be an active one due to its feedback. The external signal coming into the receiver is actually the one generated by the emitter, which serves as a continuously updated initial condition for the passive dynamical process of the receiver. If the elements in the receiver are exactly the same as the ones in the emitter, it is then clear that with the same initial conditions (continuously updated at the receiver), both should produce the same signal. More rigorously, the

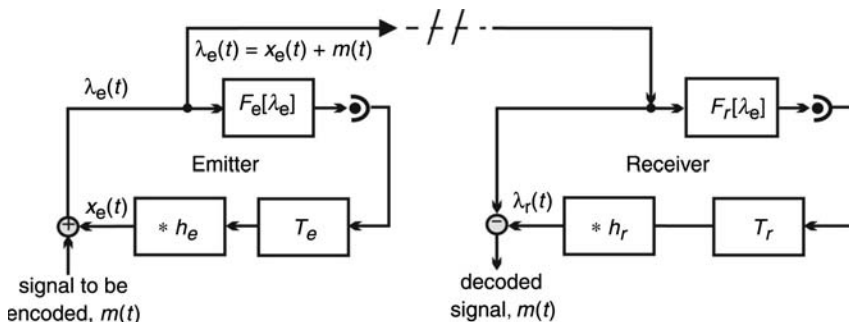


FIGURE 12.10

Bloc diagram for chaos replication and decoding.

signal generated by the receiver can be written in an integral form using the impulse response of the linear filter:

$$\lambda_e(t) = \{[h_e]_\theta \star [F_e(\lambda_e)]_{\theta-T_e}\}(t) = \int h_e(t-\theta) \cdot F_e[\lambda_e(\theta - T_e)] d\theta \quad (12.4)$$

Following the same idea and according to the description of the receiver, the equations for the replicated wavelength is:

$$\lambda_r(t) = \{[h_r]_\theta \star [F_r(\lambda_e)]_{\theta-T_r}\}(t) = \int h_r(t-\theta) \cdot F_r[\lambda_e(\theta - T_r)] d\theta. \quad (12.5)$$

It then becomes obvious that $\lambda_r(t) = \lambda_e(t)$ if we have $h_r(\cdot) = h_e(\cdot)$, $F_r(\cdot) = F_e(\cdot)$ and $T_r = T_e$. The replication is then perfect and stable against small perturbations (because the receiver is passive). The time required for the replication to occur is related to the delay, but the emitter and the receiver chaos remain identically the same as long as there is no interruption on the transmission line (to maintain the same initial condition for the dynamical process).

Equation (12.4) and Equation (12.5) can also be used to evaluate the influence of the unavoidable small mismatches when experimentally adjusting the parameter of the receiver with the ones of the emitter [8]. These equations are used to compute (in dB) the relative RMS replication error with respect to the chaos RMS amplitude:

$$(\varepsilon)_{\text{dB}} = 10 \log \frac{\langle [\lambda_e(t) - \lambda_r(t)]^2 \rangle}{\langle [\lambda_e(t)]^2 \rangle} \quad (12.6)$$

This error is represented in Figure 12.11 for mismatch on the parameter β and Φ_0 [see Equation (12.2) and Equation (12.3)]. It can be seen that a

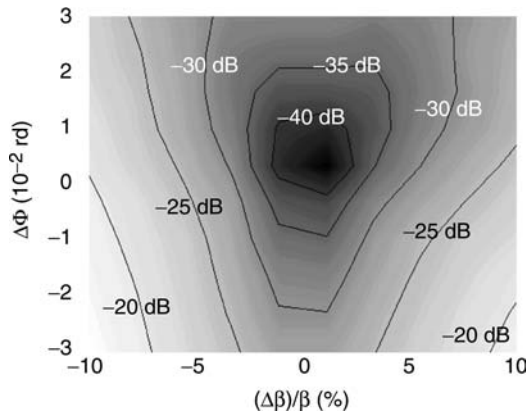


FIGURE 12.11

Replication error against parameter mismatch.

small percentage can be tolerated for a decoding noise error lower than -30 dB. However, the tolerance is not exactly the same for each parameter, especially for the delay. This fact can be explained as follows. Since we are requiring an exact temporal replication with signals varying with a typical time scale τ , the delay has to be adjusted approximately with the precision of small percentage of τ , but not of T . In the large delay case, the ratio T/τ might be rather high, and hence the tolerance for T is given by $\delta T/T \simeq (\delta\tau/\tau)/(T/\tau)$. A typical precision of 0.01% is required in the experiment for the parameter T , which is known to be the most critical one.

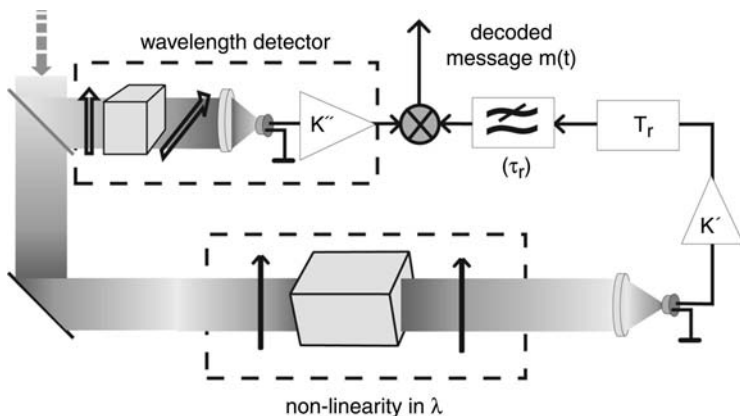
Consider now that a message is added inside the emitter oscillation loop. Using a similar way as for the replication, it can be found that the message is recovered at the receiver when we subtract from the received signal (the one that contains the message) the replicated chaos (the one generated by the receiver). The recovered message is of course superimposed on a decoding noise due to the parameter mismatch. An important consequence of the limited accuracy in the emitter–receiver matching is that the message cannot be hidden in the chaos at the emitter with an amplitude smaller than the decoding noise, otherwise the message cannot be correctly recovered. This leads to a compromise between the masking efficiency (how deep the message is masked in the chaos, the message-to-chaos ratio: $MCR < 1$) and the decoding quality (how good the message is recovered, the signal-to-noise ratio: SNR at the decoding > 1):

$$\left[\frac{\text{Chaos}}{\text{Error}} \right]_{\text{dB}} = \text{SNR}_{\text{dB}} - \text{MCR}_{\text{dB}} \quad (12.7)$$

One should try then to have as good matching as possible, but also a chaos with as much energy as possible.

12.4.3 Experimental Results

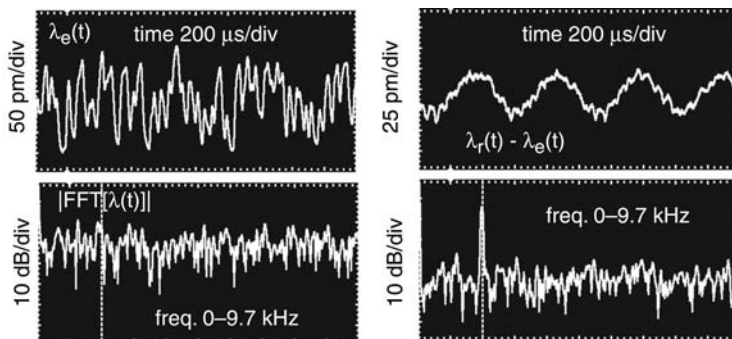
Figure 12.12 represents the wavelength chaos decoder set-up. When compared with the emitter set-up in Figure 12.7, the main difference is the absence of a tunable DBR laser at the receiver. This is actually not required, since the subtraction between the received chaos and the one locally generated is performed electronically (subtracting wavelengths is not simple!). As the DBR laser is considered as a linear element (between the input current and the output wavelength), we have designed an optoelectronic wavelength detector (the reciprocal function of the DBR laser), using the linear part of a spectral filter. Wavelength fluctuations are thus converted linearly into optical intensity fluctuations, the latter being detected by a photodiode; an electrical current proportional to the input beam wavelength is then generated. This electrical signal enters the negative input

**FIGURE 12.12**

Set-up of the wavelength chaos receiver–decoder.

of the subtractor, the chaos replicated by the receiver being the positive input (replication is processed via the duplicated emitter elements, the wavelength nonlinearity, the detector, the time delay, and the filter).

Figure 12.13 corresponds to experimental traces when a sine waveform is encoded inside the chaos. The upper traces are in the time domain, and the lower ones are the corresponding spectra. The left traces represent the signals available on the transmission line, they exhibit the characteristics of a noise (wide and flat spectrum); a cursor indicates the position of the frequency in the spectrum which is properly hidden in the chaos. The traces on the right side are the decoded signals at the emitter. The sine waveform is clearly recovered in the time domain as well as in the frequency domain. The spectrum also allows to measure the background noise level

**FIGURE 12.13**

Experimental traces while encoding and decoding a sine waveform.

which is due to the parameter mismatch. An SNR of about 25–30 dB can be measured, but recent advances with another optoelectronic set-up allowed to match more accurately the different parameters, resulting in an SNR of nearly 40 dB.

12.5 Confidentiality and Cryptanalysis Approach

12.5.1 Key-Space Dimension

According to the previous principles, the most important thing to know for the decoding is the values of the different parameters p_i involved in the chaos generation process. Those values do not actually need to be exactly known, only a finite precision δp_i is required.

Assuming an exhaustive search of each parameter value p_i with a corresponding precision δp_i , the number of bit equivalents to the encryption key is deduced from: $N_b = \sum \log_2[p_i/\delta p_i]$. A typical number of bits for the system described by Equation (12.2) and Equation (12.3) would give approximately 50, which is a relatively small number compared with the usual ones in the algorithm-based technique (key of 256 bits length [9]). However, the encryption principle does not imply a fixed nonlinear function profile, neither theoretically nor experimentally. As already pointed out, any optical filter profile exhibiting at least one extrema is enough to generate a high complexity chaotic carrier. Slightly more complex chaos generator architecture can be considered, involving several feedback loops with different nonlinear functions, different delays, and different dynamical processes. Taking all this into account would easily increase the key length above 1000.

12.5.2 Dynamics Identification from Time Series

Similarly to algorithm-based encryption systems, the exhaustive search is usually not the optimal one for Eve. She may find other analysis techniques using the so-called “cypher text” (the signal available on the transmission line), to retrieve those parameters. Some particular equations (e.g., the Mackey–Glass [10]) of the general form (12.2) have already been extensively studied from the time series point of view [11, 12]. The consequence is that it is possible to analyze a time series generated by Equation (12.2) with a computer nowadays, and find out in a few minutes all the relevant parameters in the determinism of the system [13]. However, the involved techniques are quite complicated; they seem to be limited to single extrema nonlinear functions, and they do not seem to be robust enough

when analyzing the real experimental signal (they are usually tested with numerically generated chaos). These methods also do not seem to be applicable with more complicated architecture, as the ones already noticed in the previous section (higher order, several feedback loops, etc.).

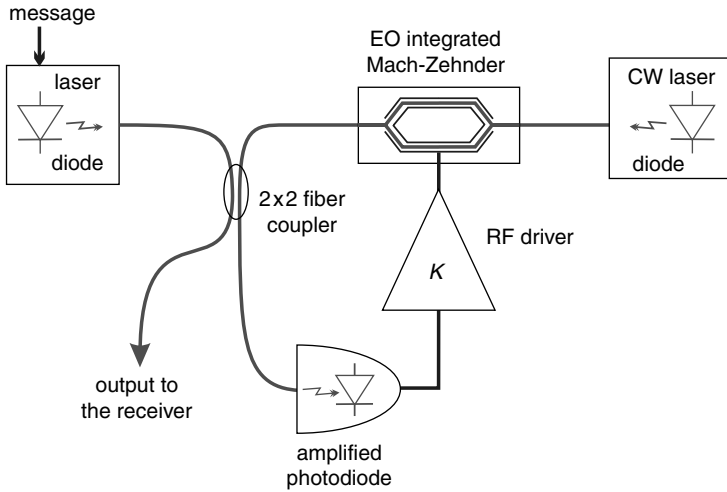
12.6 Other Optical Set-Ups for Encryption

In contrast to quantum cryptography, chaos-based cryptography is not claim absolute security, but it is more attached to develop an alternate encryption scheme for high bandwidth telecommunication systems. Its hardware-based architecture allows real-time decoding procedure, without the need for any computation time as is the case for algorithm-based encryption. Since optical transmission systems are the ones with the widest bandwidth, many different systems are currently being developed for fiber telecommunication networks. The first demonstration in optics using the previously described wavelength chaos system was only able to encode voice data, but rapidly other systems were carried out, based on fiber laser or semiconductor laser or electro-optic modulator, and reaching a few hundreds of megabits per second [14, 15] and even the gigabits per second [16]. The following sections are devoted to the brief description of other set-ups that are still studied in the field of encryption using chaos for optical telecommunications.

12.6.1 Chaotic Laser Intensity Using Electro-Optic Devices

This set-up is mainly explored by our group at the optics Dept/FEMTO-ST to demonstrate encryption speed higher than the gigabits per second in the frame of the OCCULT (optical chaos communication using laser transmitter) project, an European IST FET program. The basic principles are those of the wavelength chaos generator, in which the wavelength tunable followed by the optical spectral filter element is replaced by an ultrafast integrated electro-optic Mach–Zehnder modulator, a standard component in the present-day ultrafast optical telecommunication systems (bandwidth >10 GHz).

The experimental setup is depicted in Figure 12.14. Following the description given in the general physical principles, the linear amplifying element to be identified consists of a fast photodiode, followed by an RF amplifier, an RF driver, and the electrodes of the integrated electro-optic modulator. The linear relationship is thus calculated between the input intensity detected by the photodiode and the refractive index change in the optical waveguides due to the electro-optic effect. Since the Mach–Zehnder

**FIGURE 12.14**

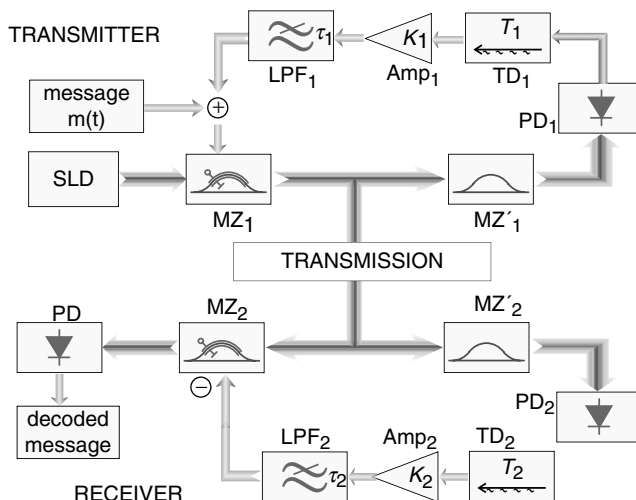
Electrooptic intensity chaos emitter.

is a two-wave interferometer, we find once again a sine square function as the nonlinear function, between the input refractive index modulation and the output intensity (the Mach-Zehnder being illuminated by a constant optical power, the CW laser diode). The message is inserted all-optically using a conventional 2×2 fiber coupler, thus leading to two outputs: one for the internal delayed feedback (the delay is performed by the fibered optical feedback length) and the other for the signal to be transmitted.

The main characteristics of this set-up consist of the use of standard or nearly standard component of optical telecommunication, all of them being connected with fibers, thus improving the stability of the set-up. These components are commercially available, they can operate at very high bit rate, and in that sense, they seem to be very promising for ultrafast chaos encryption.

12.6.2 Chaotically Coherence Modulated Light Beam

This exotic set-up [17] uses components very similar to the previous one (see Figure 12.15), but it operates physically very differently, since it involves an unusual modulation technique in optics, called the coherence modulation. The main difference in this set-up concerns the use of a broad band optical source [super luminescent diode (SLD), instead of a laser] and an unbalanced electro-optic Mach-Zehnder. We refer the reader to the references for more information, since details regarding the coherence modulation would be out of the scope of this chapter.

**FIGURE 12.15**

Emitter–receiver set-up using chaos in coherence modulation.

However, it has very attracting features [18], although it cannot use standard optical telecommunication components (the unbalanced EO Mach–Zehnder modulators). Owing to the coherence modulation properties it offers a second level of encryption, thus allowing a record masking efficiency down to -70 dB. The transmitted light does not indeed exhibit any intensity modulation, thus performing the second encryption level since a simple photodiode cannot extract any information. The message extraction can be performed all-optically, whereas all the schemes reported up to now require an electronic subtractor. Finally, the intrinsic properties allow the use of multiplexed encryption signals on the same fiber (this feature is similar to the one concerning the possibility of having multiple users in the CDMA modulation technique).

12.6.3 “All-Optical” External Cavity Laser Diode

This set-up uses a well-known and intensively studied configuration in the semiconductor laser sciences: the external cavity laser diode (ECLD, see Figure 12.16). A laser diode is very sensitive to any optical feedback, even with a very small amount of re-injected light (down to -70 dB). The basic equations describing the dynamics are known for more than 20 years [19], but the numerous behaviors are so complex and so different that the system is still being studied. The dynamics is described by a set of three first-order coupled differential equations (the rate equations of the laser), ruling the amplitude and the phase of the electric field, and the density of

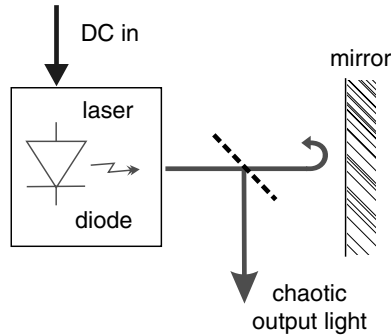


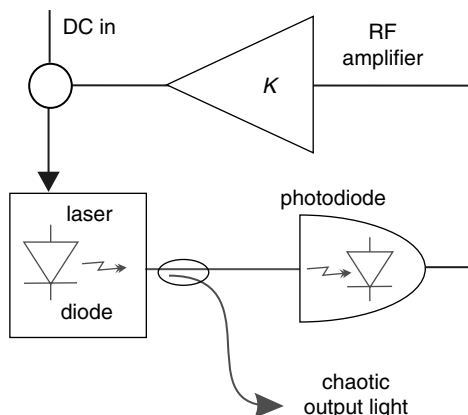
FIGURE 12.16
Basic ECLD set-up.

the inversion population producing the optical amplification in the semiconductor junction. The population equation is nonlinearly coupled to the field energy (through the gain saturation), thus resulting in a nonlinear filtering process. The delay term is introduced by the feedback field coming from the external cavity, but this delayed feedback is purely linear. The relevant parameters determining the actual dynamical regime are the feedback level, the pumping rate, and the length of the external cavity.

The first numerical demonstration of the potential use of this configuration [20] for chaotic encryption was proposed in 1996. But the experimental demonstration of the possible synchronization between two such distant ECLDs using a unidirectional coupling appeared only a few years later [21]. This architecture is currently explored in the OCCULT project by five different teams in Europe. Its main advantage lies on a very fast dynamical process, with time scales as short as a few picoseconds (related to the carrier lifetime in the active junction and to the photon life time in the free running laser cavity), but it might suffer from an extreme experimental sensitivity.

12.6.4 Direct Optoelectronic Feedback in SC Lasers

Although it appears as the most simple configuration, it owns the record speed for chaos encrypted binary data, with 2.5 Gbits/sec [22]. The chaos generator is depicted in Figure 12.17. It involves both electronic and optoelectronic dynamical systems. It consists of a high-speed semiconductor laser diode whose injection current is driven from the direct, but delayed and eventually amplified, detection of the output optical power by a fast photodiode. The dynamical process once again involves the rate equations of a semiconductor laser, but the delayed term influences the rate of the population inversion (instead of the electric field) through the feedback

**FIGURE 12.17**

Direct optoelectronic feedback in an SC laser.

current, and its influence is nonlinearly converted by the feedback path (from the photodiode to the laser junction). The system has been shown to be very fast. The drawbacks concern the low dimension of the chaos produced (Lyapunov dimension <10 [23]) and a low flexibility of the set-up (low key-space dimension).

12.7 Conclusions

Although many samples have appeared in the last 10 years, chaos-based encryption systems have to be considered as still being under development, and important goals still need to be achieved to improve their efficiency in terms of noise sensitivity, decoding quality, robustness, and confidentiality. More specific efforts have to be taken regarding the modulation technique (i.e., the way used to mix the information into the chaotic waveform). Also, the actual and precise evaluation of the degree of confidentiality is currently being studied. Even if some of the proposed encryption schemes have been broken [24], the field is quite young and the evolution possibilities are enormous.

Chaos cryptography is not necessarily a competitor for classical algorithm-based encryption techniques; in contrast, it might give ideas to develop new algorithms. It is, for example, well known that computer science already uses a random number generator algorithm based on chaotic dynamics; an even more relevant example is a recent encryption software commercialized in Japan, which is based on chaos theory, and which is claiming a 2048 bit encryption key.

References

1. C. Mira, S. Gentil, R. Abraham, H. Kawakami, M. Hasler, C. Grebogi, L. Chua, Y. Kevrekidis, and Y. Maistrenko, Bifurcation, chaos, transformations non inversibles, applications, École d'Automatique d'Été de Grenoble, Laboratoire d'Automatique de Grenoble, France, September 1995.
2. K. Ikeda, Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system, *Optics Comm.*, 30 (2), 1979.
3. J. Duvernoy, J.P. Goedgebuer, and H. Porte, Bistabilité, multistabilité et chaos en longueur d'onde, *Ann. des Télécommun.*, 42 (5-6), 315, 1987.
4. L. Larger, J.-P. Goedgebuer, and J.M. Merolla, Chaotic oscillator in wavelength: a new set-up for investigating differential difference equations describing non linear dynamics, *IEEE J. Quantum Electron.*, 34 (4), 594–601, 1998.
5. L.M. Pecora and T.L. Carroll, Synchronization in Chaotic Systems, *Phys. Rev. Lett.*, 64 (8), 1990.
6. H. Dedieu, M.P. Kennedy, and M. Hasler, Chaos Shift Keying: modulation and demodulation of a chaotic carrier using self-synchronizing Chua's circuits, *IEEE Trans. Circuits Syst.*, 40 (2), 1993.
7. J.-P. Goedgebuer, L. Larger, and H. Porte, An optical cryptosystem based on replication of hyperchaos and wavelength-induced nonlinearities, *Phys. Rev. Lett.*, 80 (10), 2249–2252, 1998.
8. L. Larger, J.-P. Goedgebuer, and F. Delorme, An optical encryption system using hyperchaos generated by an optoelectronic wavelength oscillator, *Phys. Rev. E*, 57 (6), 6618–6624, 1998.
9. S. Singh, *Histoire des Codes Secrets*, J.C. Lattès, 1999.
10. M.C. Mackey and L. Glass, *Science*, 197, 287, 1977.
11. R. Hegger, M.J. Bünner, and H. Kantz, Identifying and modeling delay feedback systems, *Phys. Rev. Lett.*, 81 (3), 558–561, 1998.
12. C. Zhou and C.H. Lai, Extracting messages masked by chaotic signals of time-delay systems, *Phys. Rev. E*, 60 (1), 320–323, 1999.
13. V.S. Udaltsov, L. Larger, J.-P. Goedgebuer, J.-B. Cuenot, P. Levy, and W.T. Rhodes, Cracking chaos-based encryption systems ruled by nonlinear time delay differential equations, *Phys. Lett. A*, 308, 54–60, 2003.
14. G.D. VanWiggeren and R. Roy, Communicating with chaotic lasers, *Science*, 279 (3), 1198–1200, 1998.
15. J.-P. Goedgebuer, P. Levy, L. Larger, C.C. Chen, and W.T. Rhodes, High bandwidth chaotic encryption system, *IEEE J. Quant. Electron.* (Special Issue on Optical Chaotic Cryptography), 38 (9), 1178–1183, 2002.
16. J. Paul, S. Sivaprakasam, P.S. Spencer, P. Rees, and K.A. Shore, GHz bandwidth message transmission using chaotic diode lasers, *Electron. Lett.*, 38 (1), 28–29, 2002.
17. L. Larger, M.W. Lee, J.-P. Goedgebuer, T. Erneux, and W. Elflein, Chaos in coherence modulation: bifurcations of an oscillator generating optical delay fluctuations, *JOSA B*, 18 (8), 1063–1068, 2001.
18. M.W. Lee, L. Larger, and J.-P. Goedgebuer, Encryption system using chaotic delays between lightwaves, *IEEE J. Quantum Electron.*, 39 (7), 931–935, 2003.

19. R. Lang and K. Kobayashi, External optical feedback effects on semiconductor injection laser properties, *IEEE J. Quantum Electron.*, 16 (3), 347–354, 1980.
20. C.R. Mirasso, P. Colet, and P. Garcia-Fernandez, Synchronization of chaotic semiconductor lasers: application to encoded communications, *IEEE Photon. Tech. Lett.*, 8 (2), 299–301, 1996.
21. I. Fischer, Y. Liu, and P.D. Peter, Synchronization of chaotic semiconductor laser dynamics on subnanosecond time scales and its potential for chaos communication, *Phys. Rev. A*, 62 (1), 011801, 2000.
22. S. Tang and J.M. Liu, Message encoding–decoding at 2.5 Gb/s through synchronization of chaotic pulsing semiconductor lasers, *Optics Lett.*, 26 (23), 1843–1845, 2001.
23. H.D.I. Abarbanel, M.B. Kennel, L. Illing, S. Tang, H.F. Chen, and J.M. Liu, Synchronization and Communication Using Semiconductor Lasers With Optoelectronic Feedback, *IEEE J. Quantum Electron.*, 37 (10), 1301–1311, 2001.
24. J.B. Geddes, K.M. Short, and K. Black, Extraction of signals from chaotic laser data, *Phys. Rev. Lett.*, 83 (25), 5389–5392, 1999.

13

Indirect Field-Oriented Control of Induction Motors: A Hopf Bifurcation Analysis

Francisco Gordillo, Francisco Salas, Romeo Ortega, and Javier Aracil

CONTENTS

13.1 Introduction	481
13.2 The Hopf Bifurcation and the Emergence of Oscillations	483
13.3 Detection of Hopf Bifurcations in Induction Motors with Indirect FOC	485
13.3.1 Zero Load Torque Case	486
13.3.2 Nonzero Load Torque Case	488
13.4 Simulations	491
13.5 Approximate Study of Limit Cycles Using the Harmonic Balance Method	493
13.5.1 Zero Load Torque Case ($\tau_L = 0$)	495
13.5.2 Nonzero Load Torque Case ($\tau_L \neq 0$)	496
13.6 Conclusions	497
Appendix: Motor Model	497
References	501

13.1 Introduction

Field-oriented control (FOC) is the standard for high dynamic performance induction motor drives due to its high reliability. Historically, this remarkable controller was derived as a result of physical intuition and a deep understanding of the machine operation, with little concern about a rigorous analytical study of its stability and performance. An approximate

analysis (based on steady-state behavior, time-scale assumptions, and linearizations [1–3]) can be combined with the designer expertise to commission the controller in simple applications. However, to meet large bandwidth requirements, or other tight specifications, this *ad hoc* commissioning stage may be time-consuming and expensive. To simplify the off-line tuning of FOC, and eventually come to terms with its achievable performance, a better theoretical understanding of the dynamic behavior of FOC is essential. Such an analysis is unfortunately stymied by the fact that the dynamic behavior of the closed loop is described by complex nonlinear relationships.

Realizing the practical importance of FOC, and motivated by the need to clarify its theoretical underpinnings, a series of studies have been carried out in the last few years on indirect FOC for current-fed induction machines. The outcome of this research is summarized in Ref. [4]. In this chapter, we continue with this line of research and concentrate on the practically important problem of *oscillation quenching* through suitable tuning of the gains of the PI velocity loop. The appearance of self-sustained oscillations in high-performance AC drives, and particularly in FOC of induction motors, is a well-documented, but little understood phenomenon. It is well known that the oscillations may be quenched by retuning the outer-loop PI speed control, but no precise rules to carry out this task are known. This tuning procedure is particularly difficult due to the high uncertainty on the rotor time-constant. In this chapter, we apply some standard techniques of dynamical systems (in particular, harmonic balance and bifurcation analysis) to show that these oscillations may arise due to the existence of Hopf bifurcations. As a result of our analysis, we obtain some simple rules to quench the oscillations via a suitable PI tuning.

The present study was motivated by Ref. [5] where the robustness results of FOC reported in Ref. [6] are generalized. In particular, in Ref. [5], an explicit construction of a Lyapunov function, which allows for the evaluation of stability margins, is given. Furthermore, the authors also study the existence of a saddle-node bifurcation. Our study is complementary to the latter as we are interested here in Hopf bifurcations, which as our study conclusively proves, are at the core of the oscillation phenomena often observed in practice [2]. Reference [7] deals with the same problem but with a different and complementary approach.

This chapter is organized as follows: in Section 13.2, we briefly review the theory of Hopf bifurcations. In Section 13.3, the bifurcation is detected with standard tangent approximation analysis. Simulations that corroborate the theoretical predictions are given in Section 13.4. Section 13.5 is devoted to a complementary analysis of bifurcation detection via harmonic balance. We end the chapter with Section 13.6, which contains some concluding remarks. The derivation of the system equations is given in the Appendix.

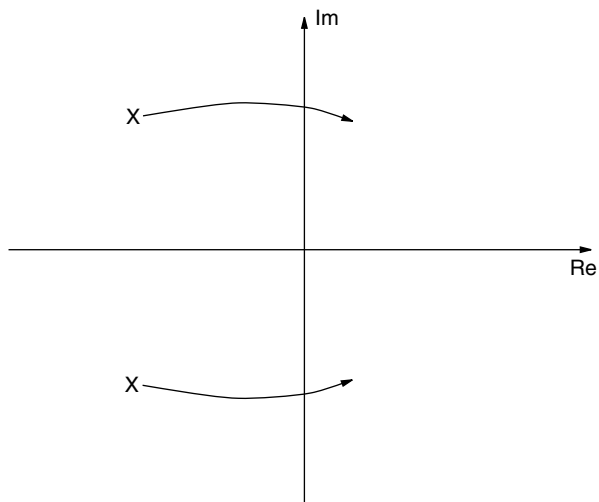
13.2 The Hopf Bifurcation and the Emergence of Oscillations

Bifurcation analysis in control systems is a growing field [8–11]. This is mainly due to the realization of the richness and variety of behaviors which a nonlinear system can display [12]. These behaviors include self-sustained oscillations, which are out of the scope of the realm of linear systems [13, 14]. The most frequent bifurcation associated with the appearance of oscillations is the Hopf bifurcation, which will be recalled later in this section.

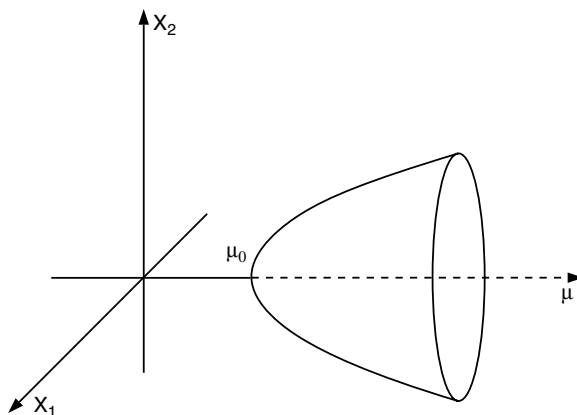
The analysis of a control system has been mainly concerned with the analysis of the system around the operating point. This point is the desired attractor of the closed loop system. The stability of the corresponding equilibrium point can be analyzed by linearizing the system around that point. The eigenvalues of the linearized system define the stability of the system. The stability is lost when the real part of one or more of the eigenvalues becomes positive. Generically, this loss of stability is reached when one real eigenvalue becomes positive or when two complex eigenvalues cross the imaginary axis. In this last case, the phenomenon known as Hopf bifurcation is produced.

Suppose that the dynamical system $\dot{x} = f(x, \mu)$ with $x \in \mathbb{R}^n$ and $\mu \in \mathbb{R}$ has an equilibrium point at x^0 , for some $\mu = \mu^0$; that is, $f(x^0, \mu^0) = 0$. Let $A(\mu) = D_x f(x^0(\mu), \mu)$ be the Jacobian matrix of the system at the equilibrium point. Assume that $A(\mu^0)$ has a single pair of purely imaginary eigenvalues $\lambda(\mu^0) = \pm j\omega$, and no other eigenvalue with zero real part; and furthermore, $d\Re(\lambda(\mu))/d\mu|_{\mu=\mu^0} \neq 0$. The last condition is known as the transversality hypothesis. In the complex plane, the transversality hypothesis can be easily checked by the well-known root locus of the closed loop system (Figure 13.1). Under these conditions, the Hopf bifurcation theorem states that a limit cycle is born at (x^0, μ^0) [15]. The parameter μ is known as the bifurcation parameter. Actually, there is a third condition for the existence of Hopf bifurcations, but this is a technical hypothesis in order to avoid nongeneric, degenerate cases such as the linear one [14].

The Hopf bifurcation can be visualized as shown in Figure 13.2. As the Hopf bifurcation is produced, the operating point loses the stability, and a stable limit cycle appears around it. This is known as the supercritical Hopf bifurcation. There is also a subcritical one, where an unstable limit cycle dies at the bifurcation point. The Hopf bifurcation that will be detected in the sequel in the induction motor is of the supercritical type, that is, the stable limit cycle surrounds the unstable equilibrium state.

**FIGURE 13.1**

Transversality condition for a Hopf bifurcation.

**FIGURE 13.2**

Supercritical Hopf bifurcation.

One procedure for detecting Hopf bifurcations is the following:

1. Look for the equilibrium points of the system solving the equation $f(x, \mu) = 0$. The equilibria x^0 will depend on μ ; that is $x^0 = x^0(\mu)$.
2. For each equilibrium point:
 - (a) Obtain the value of the Jacobian matrix of the system at the equilibrium point.

- (b) Look for the values of the bifurcation parameter μ which cause the Jacobian to have two eigenvalues in the imaginary axis. This condition can be formulated for any dimension n in the following way [16]:

$$H_{n-1}(\mu) = 0 \quad (13.1)$$

$$H_{n-2}(\mu) > 0, H_{n-3} > 0, \dots, H_1(\mu), p_0(\mu) > 0 \quad (13.2)$$

where H_i stands for the i principal minor of the Hurwitz matrix of the characteristic polynomial of the Jacobian matrix and p_0 is the zero-order term of this polynomial.

If a Hopf bifurcation occurs, then there will exist values of μ which fulfill Equation (13.1) and Equation (13.2). This condition is necessary but not sufficient. In order to ensure the emergence of a limit cycle, the eigenvalues must cross the imaginary axis when varying the bifurcation parameter μ (transversality condition, see Figure 13.1). This condition can be checked by plotting the root locus of the linearized system with μ as a parameter.

It should be noted that the bifurcations give the conditions for the loss of stability. But bifurcations not only indicate that the stability is lost, but also *how* this loss of stability is produced. In the case of the supercritical Hopf bifurcation, the emergence of a stable limit cycle is associated with the loss of stability, and the system oscillates. Therefore, in looking for bifurcations, more information about the dynamical phenomena to be expected at the boundaries of the stability region is obtained.

The conditions for the emergence of a Hopf bifurcation only predict when an oscillation appears. In the general case, the study of that oscillation is a very difficult task, but an approximation to it can be reached by means of the harmonic balance method [17, 18]. For a mathematical analysis of the validity of the harmonic balance approximation, see Refs. [19, 20].

13.3 Detection of Hopf Bifurcations in Induction Motors with Indirect FOC

In this section, the aforementioned procedure is applied to detect Hopf bifurcations in a current driven induction motor with indirect FOC. The

equations of the system, which are developed in the Appendix, are

$$\dot{x}_1 = -c_1 x_1 + \frac{\kappa c_1}{u_2^0} x_2 x_4 + c_2 u_2^0 \quad (13.3)$$

$$\dot{x}_2 = -c_1 x_2 - \frac{\kappa c_1}{u_2^0} x_1 x_4 + c_2 x_4 \quad (13.4)$$

$$\dot{x}_3 = -c_4 [c_5 (x_1 x_4 - x_2 u_2^0) - \tau_L] \quad (13.5)$$

$$\dot{x}_4 = k_i x_3 - k_p c_4 [c_5 (x_1 x_4 - x_2 u_2^0) - \tau_L] \quad (13.6)$$

The notation is explained in the Appendix. Assuming that the rotor inductance is known, the parameter κ is equal to the ratio between the estimated and the true values of the rotor resistance, that is, $\kappa = \hat{R}_r / R_r$. The rotor resistance is really unknown due to the fact that it changes with time. Therefore, κ has been chosen as the bifurcation parameter. Notice that only positive values of κ are of physical interest. The analysis in this paper is directed towards the causes of the loss of stability. Thus, the desired equilibrium point is assumed to be locally stable in the case of exact estimation of the rotor resistance $\kappa = 1$. The aim of the study is to find the range of values for κ , around $\kappa = 1$, for which the equilibrium preserves its stability, by looking for the emergence of bifurcations. It must be pointed out that other complex phenomena due to the emergence of more bifurcations may occur beyond that range but these phenomena are not the subject of this paper. Two cases will be examined here: the case when there is no load torque ($\tau_L = 0$) and the general case with $\tau_L \neq 0$.

13.3.1 Zero Load Torque Case

When $\tau_L = 0$, the analysis is simpler because in this case the system has only one equilibrium point which is independent of the rest of parameters [6]. Indeed, the equilibrium is $x^0 = (x_1^0, x_2^0, x_3^0, x_4^0) = (\frac{c_2}{c_1} u_2^0, 0, 0, 0)$ and can be obtained by equating the time derivatives of x_i to zero in system (13.3) to system (13.6) and solving the resulting system of equations.

The Jacobian matrix of system (13.3) to system (13.6) at the equilibrium point has the expression

$$A(x^0) = \begin{bmatrix} -c_1 & 0 & 0 & 0 \\ 0 & -c_1 & 0 & c_2 - \kappa c_2 \\ 0 & c_4 c_5 u_2^0 & 0 & -\frac{u_2^0 c_4 c_5 c_2}{c_1} \\ 0 & k_p c_4 c_5 u_2^0 & k_i & -\frac{u_2^0 k_p c_4 c_5 c_2}{c_1} \end{bmatrix} \quad (13.7)$$

Looking at the first row of the Jacobian matrix, it is obvious that $\det(c_1 I - A(x^0)) = 0$. Therefore, $-c_1$ is an eigenvalue of $A(x^0)$ and, then, its characteristic polynomial must be divisible by $\lambda + c_1$. Indeed

$$\det(\lambda I - A(x^0)) = (\lambda + c_1)(\lambda^3 + \alpha_2 \lambda^2 + \alpha_1 \lambda + \alpha_0)$$

with

$$\alpha_2 = \frac{c_1^2 + u_2^0 k_p c_4 c_5 c_2}{c_1} \quad (13.8)$$

$$\alpha_1 = \frac{u_2^0 k_p c_4 c_5 c_1 \kappa c_2 + c_4 c_5 k_i c_2 u_2^0}{c_1} \quad (13.9)$$

$$\alpha_0 = u_2^0 c_4 c_5 k_i \kappa c_2 \quad (13.10)$$

Condition (13.1) and Condition (13.2) for the particular case of $n = 3$ can be written as

$$\alpha_0 = \alpha_2 \alpha_1 \quad (13.11)$$

$$\alpha_1 > 0 \quad (13.12)$$

$$\alpha_0 > 0 \quad (13.13)$$

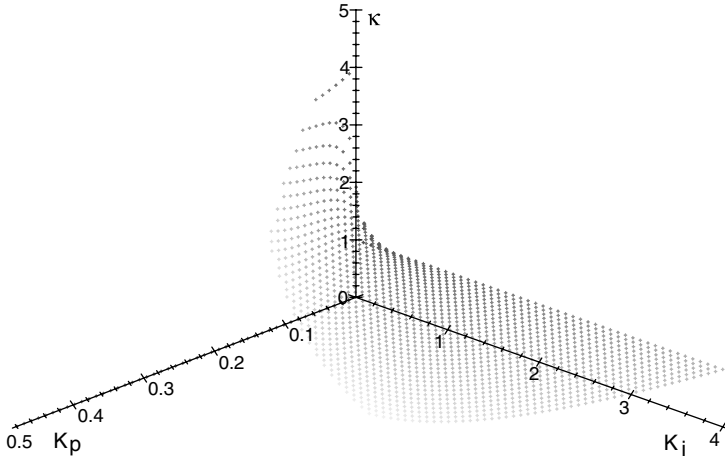
Since all the parameters in Equation (13.9) and Equation (13.10) are positive, Equation (13.12) and Equation (13.13) hold. Equation (13.11) leads to

$$\frac{(c_1^2 + u_2^0 k_p c_4 c_5 c_2)(u_2^0 k_p c_4 c_5 c_1 \kappa c_2 + c_4 c_5 k_i c_2 u_2^0)}{c_1^2} = u_2^0 c_4 c_5 k_i \kappa c_2$$

which yields

$$\kappa = -\frac{(c_1^2 + u_2^0 k_p c_4 c_5 c_2) k_i}{c_1 (k_p c_1^2 + u_2^0 k_p^2 c_4 c_5 c_2 - c_1 k_i)} \quad (13.14)$$

This expression is of great practical interest. It has been assumed that the desired equilibrium is stable for $\kappa = 1$. In that case, no limit cycles will appear in the neighborhood of this point provided that this equilibrium does not suffer a Hopf bifurcation. Expression (13.14) gives the value of $\kappa = \hat{R}_r / R_r$ corresponding to a Hopf bifurcation in the case of $\tau_L = 0$. In other words, Equation (13.14) gives an upper limit on the error of the estimation of R_r when $\tau_L = 0$ in order to prevent the existence of limit cycles. In Ref. [6], the authors prove that for $\kappa < 3$, there is only one equilibrium point,

**FIGURE 13.3**

Representation of Equation (13.14) for $c_1 = 4$, $c_2 = 4$, $c_4 = 1$, $c_5 = 1$, and $u_2^0 = 1$.

and necessary and sufficient conditions are given to assure the stability of this equilibrium. Here, we look at the cause of the loss of stability: the emergence of a limit cycle by a Hopf bifurcation.

In Figure 13.3, the values of κ given by Equation (13.14) are plotted against k_p and k_i for $c_1 = 4$, $c_2 = 4$, $c_4 = 1$, $c_5 = 1$, and $u_2^0 = 1$. For each pair of values of k_p and k_i , the system will not present limit cycles provided that κ is below the curve. As can be seen from Equation (13.14), the curve corresponding to the aforementioned values goes to infinity for $k_i = k_p^2 + 4k_p$. For $k_i < k_p^2 + 4k_p$, the critical value of κ is less than zero and has no practical meaning.

A practical conclusion from Equation (13.14) and Figure 13.3 is that increasing k_p allows for larger values of the critical κ (larger admissible estimation errors), and increasing k_i yields the opposite effect.

13.3.2 Nonzero Load Torque Case

The case with $\tau_L \neq 0$ is more involved and a single expression similar to Equation (13.14) is difficult to obtain. The reason is that, for this case, the equilibrium point is obtained by solving a third-order polynomial equation [6]. Substitution of the value of the equilibrium in the Jacobian matrix leads to complex expressions. Nevertheless, for this case conclusions can be drawn for concrete values of the parameters. In fact, for particular values of

$c_1, c_2, c_4, c_5, u_2^0, k_p, k_i$, and τ_L , the critical value of κ may be obtained following the same procedure as mentioned earlier. For example, making $c_1 = 4, c_2 = 4, c_4 = 1, c_5 = 1, u_2^0 = 1, k_p = 0.1, k_i = 1$, and $\tau_L = 0.3$, the eigenvalues of the Jacobian matrix at the corresponding equilibrium point can be computed for each value of κ , giving the root locus of Figure 13.4 (in this figure, only two branches are plotted because the other two branches are far away to the left). It can be seen in this figure that the condition of transversality is fulfilled, and a Hopf bifurcation is expected to be produced.

Figure 13.5 shows the values of the critical κ for several values of τ_L , and k_p (the rest of parameters have the same values as shown).

The most critical value of κ corresponding to the Hopf bifurcation will be that which is closest to one since it has been assumed that the desired equilibrium is stable for $\kappa = 1$. As can be seen in Figure 13.5, this case corresponds to $\tau_L = 0$. If this fact were general, it would give more practical value to Equation (13.14), but is difficult to be proven for the general case. Nevertheless, it seems to be true in accordance with all the cases tested.

Figure 13.5 also shows that as τ_L increases so does the value of κ corresponding to a Hopf bifurcation. It could be thought that, for high values

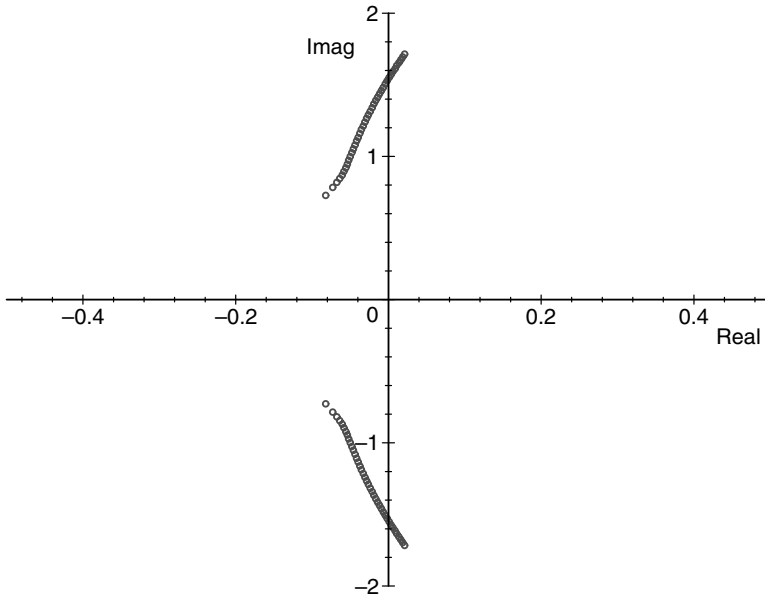
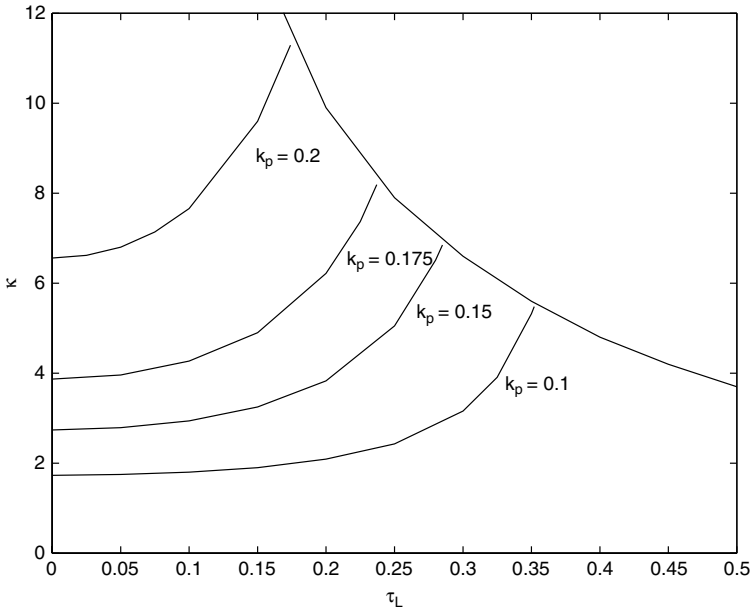


FIGURE 13.4

Root locus for $c_1 = 4, c_2 = 4, c_4 = 1, c_5 = 1, u_2^0 = 1, k_p = 0.1, k_i = 1$ and $\tau_L = 0.3$.

**FIGURE 13.5**

Values of κ corresponding to a Hopf bifurcation vs. τ_L .

of τ_L , the estimation error would not affect the stability of the system. Nevertheless, this stability is lost in another bifurcation, a saddle-node bifurcation, due to the emergence of two new equilibria, as has been reported in Ref. [5, 6]. The upper curve in Figure 13.5 represents the values of κ corresponding to this bifurcation. It must be pointed out that the Hopf bifurcation exists above the curve of the saddle-node bifurcation but it has not been plotted because it has no practical meaning because the local stability of the equilibrium has already been lost at this point. Furthermore, Figure 13.5 does not represent the full bifurcation diagram of the system, but only the curves associated with both bifurcations. Specifically, around the intersection of the two kind of curves in Figure 13.5 more complicated bifurcations will probably exist. The Hopf bifurcation studied in this paper is of a more generic character than the possible bifurcations associated with the intersection of those curves.

The practical application of this graph is the following: for a given set of parameters, the graph can be constructed. The admissible values of κ (so the equilibrium point is stable) as a function of τ_L are under two curves: the one corresponding to the given value of k_p and the one corresponding to the saddle-node bifurcation. Parameter κ must be less than the minimum of these curves (for admissible τ_L). In all the situations, we

have tested this minimum corresponds to $\tau_L = 0$ if excessively large values of τ_L are not considered. The value of the critical κ for $\tau_L = 0$ is given by Equation (13.14).

Curiously, there are situations where very large estimation errors may be admissible: for small values of the load torque ($\tau_L \rightarrow 0$) and when k_p and k_i are such that Equation (13.14) yields negative values. In this case, the saddle-node bifurcation is not possible (the upper curve of Figure 13.5 goes to infinity as $\tau_L \rightarrow 0$) and the Hopf bifurcation would occur for $\kappa < 0$, which has no practical meaning.

13.4 Simulations

To corroborate the aforementioned results, the system has been simulated with the value of the parameters given in Section 3.2 (including $k_i = 1$) but with different values for k_p , τ_L , and κ which are shown in Table 13.1.

It can be seen from Table 13.1 that simulations A and B correspond to the case of $\tau_L = 0$. Introducing the values of the parameters in Equation (13.14), this gives $\kappa = 1.73$ as the limit for the emergence of a limit cycle. Therefore, simulation A must correspond to a stable behavior and simulation B must present a limit cycle. On the other hand, simulations C and D correspond to $\tau_L = 0.2 \neq 0$ and Equation (13.14) is not applicable but Figure 13.5 is. In this case, the bifurcation point is $\kappa = 3.83$ and simulation C must not present limit cycles while simulation D must.

In performing all these simulations, the predicted behaviors are corroborated as it is shown in Figure 13.6. In this figure, two graphs are plotted for each simulation. In the first one, the evolution of the amplitudes of the oscillations can be observed. It can be seen that the amplitudes of simulations A and C go to zero whereas in simulations B and D, the oscillations tend to a limit cycle. The oscillations can be better observed in the second zoom graph of each simulation.

TABLE 13.1
Values for k_p , τ_L , and κ Corresponding to the Four Simulations

Simulation	k_p	τ_L	κ	Predicted Behavior
A	0.1	0	1.65	Stable
B	0.1	0	1.8	Limit cycle
C	0.15	0.2	3.7	Stable
D	0.15	0.2	3.9	Limit cycle

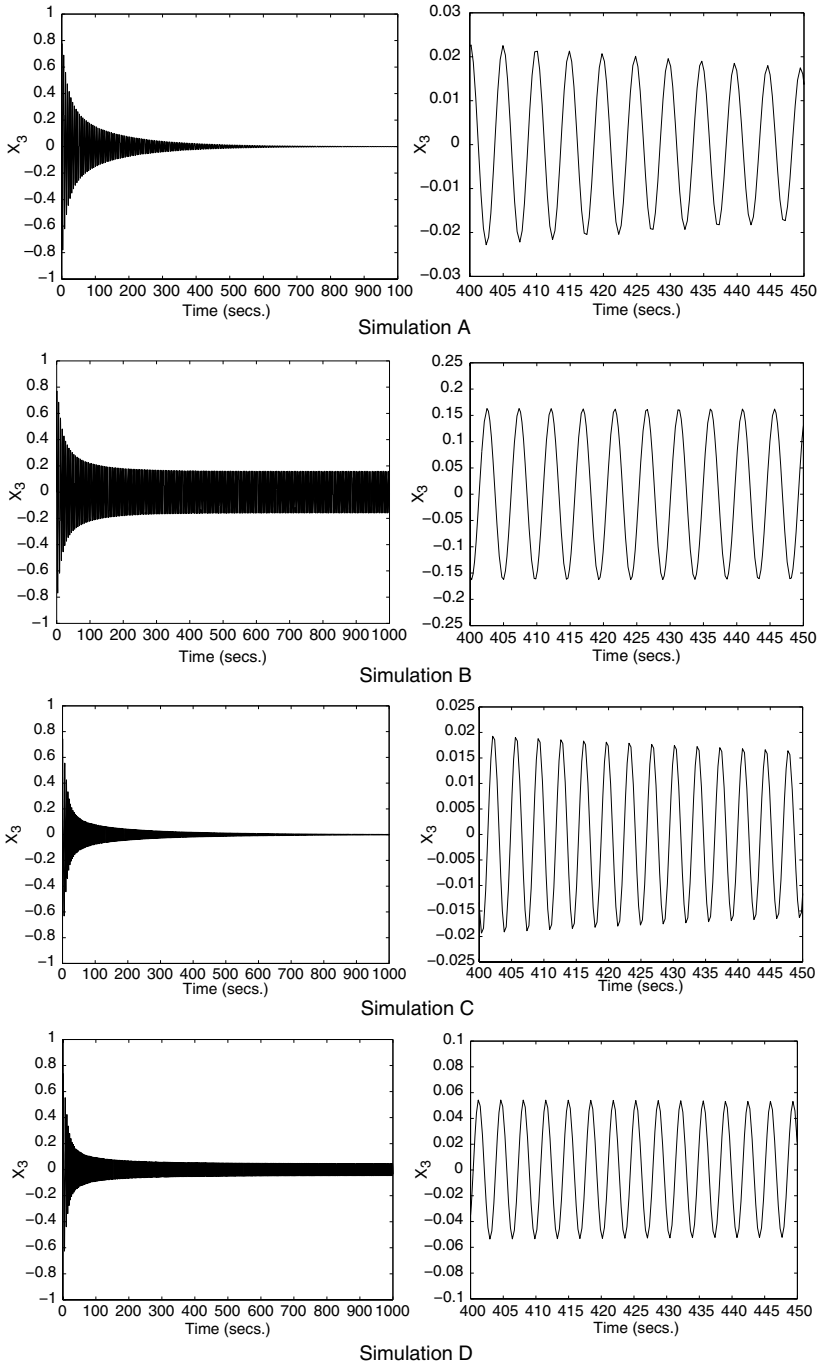


FIGURE 13.6

Evolution of x_3 in the four simulations.

13.5 Approximate Study of Limit Cycles Using the Harmonic Balance Method

Although, as stated in previous sections, the Hopf bifurcation is of a local nature, in this section, we will apply a method that is not restricted to the bifurcation point, but has a more global nature. This method, called harmonic balance, not only allows us to detect limit cycles at their birth, as in local methods, but also to evaluate an approximation to their frequency and amplitude [20].

The harmonic balance method is based on the Fourier series for periodic functions. Consider a scalar function $y(t)$, which is periodic with period T . It can be represented by its Fourier series in the form

$$y(t) = c_0 + \sum_{k=1}^{\infty} \{a_k \cos(k\omega t) + b_k \sin(k\omega t)\} \quad (13.15)$$

where ω is the angular frequency and has the value $\omega = 2\pi/T$, and c_0 is the bias coefficient.

The Bessel inequality [21] states that $a_k, b_k \rightarrow 0$ as $k \rightarrow \infty$, justifying the approximation of $y(t)$ by truncating the Fourier series after a finite number of terms,

$$y(t) = c_0 + \sum_{k=1}^n \{a_k \cos(k\omega t) + b_k \sin(k\omega t)\} \quad (13.16)$$

The simplest approximation would be to take $n = 1$, resulting in the first-harmonic balance method. In this case, every dynamical variable is approximated by a single sinusoidal oscillation plus a bias constant. This balance is equivalent to the dual describing function method [22] but, here, the harmonic balance is applied directly.

To apply the method to the fourth-order system (13.3)–(13.6), we assume that each state variable x_i has a self-sustained oscillation of the same frequency ω ; that is

$$x_1 = a_{10} + a_{11} \cos \omega t \quad (13.17)$$

$$x_2 = a_{20} + a_{21} \cos \omega t + a_{22} \sin \omega t \quad (13.18)$$

$$x_3 = a_{30} + a_{31} \cos \omega t + a_{32} \sin \omega t \quad (13.19)$$

$$x_4 = a_{40} + a_{41} \cos \omega t + a_{42} \sin \omega t \quad (13.20)$$

where the coefficients a_{i0} , ($i = 1, \dots, 4$) are the bias components and a_{i1} , a_{i2} are the terms a_1 and b_1 in Equation (13.16). Notice that the amplitude

coefficient of the sine term a_{12} in Equation (13.17) is null since the time origin can be defined arbitrarily. The rest of the sine and cosine coefficients are not null because a phase shift between the oscillations of the state variables is considered.

As stated previously, this is an approximate method because a truncated Fourier series is used (in this case, we only use the first harmonic $n = 1$), but the method can be made as exact as desired, using more terms of the Fourier series in Equation (13.17) to Equation (13.20) [15, 19, 23]. In this case, the method would be more involved.

By substituting Equation (13.17) to Equation (13.20) and its derivatives into Equation (13.3) to Equation (13.6), grouping bias, sine, and cosine terms, and ignoring harmonics higher than one, a nonlinear system of 12 equations with 12 unknown variables, which are a_{ij} and ω , can be obtained.

The system of equations is

$$0 = -c_1 a_{10} + \frac{\kappa c_1}{u_2^0} \left(a_{20} a_{40} + \frac{1}{2} a_{21} a_{41} + \frac{1}{2} a_{22} a_{42} \right) + c_2 u_2^0 \quad (13.21)$$

$$-a_{11} \omega = \frac{\kappa c_1}{u_2^0} (a_{20} a_{42} + a_{22} a_{40}) \quad (13.22)$$

$$0 = -c_1 a_{11} + \frac{\kappa c_1}{u_2^0} (a_{20} a_{41} + a_{21} a_{40}) \quad (13.23)$$

$$0 = -c_1 a_{20} - \frac{\kappa c_1}{u_2^0} \left(a_{10} a_{40} + \frac{1}{2} a_{11} a_{41} \right) + c_2 a_{40} \quad (13.24)$$

$$-a_{21} \omega = -c_1 a_{22} - \frac{\kappa c_1}{u_2^0} a_{10} a_{42} + c_2 a_{42} \quad (13.25)$$

$$a_{22} \omega = -c_1 a_{21} - \frac{\kappa c_1}{u_2^0} (a_{10} a_{41} + a_{11} a_{40}) + c_2 a_{41} \quad (13.26)$$

$$0 = -c_4 c_5 \left(a_{10} a_{40} + \frac{1}{2} a_{11} a_{41} - u_2^0 a_{20} \right) + c_4 \tau_L \quad (13.27)$$

$$-a_{31} \omega = -c_4 c_5 \left(a_{10} a_{42} - u_2^0 a_{22} \right) \quad (13.28)$$

$$a_{32} \omega = -c_4 c_5 \left(a_{10} a_{41} + a_{11} a_{40} - u_2^0 a_{21} \right) \quad (13.29)$$

$$0 = k_i a_{30} - k_p c_4 c_5 \left(a_{10} a_{40} + \frac{1}{2} a_{11} a_{41} - u_2^0 a_{20} \right) + k_p c_4 \tau_L \quad (13.30)$$

$$-a_{41} \omega = k_i a_{32} - k_p c_4 c_5 \left(a_{10} a_{42} - u_2^0 a_{22} \right) \quad (13.31)$$

$$a_{42} \omega = k_i a_{31} - k_p c_4 c_5 \left(a_{10} a_{41} + a_{11} a_{40} - u_2^0 a_{21} \right) \quad (13.32)$$

where Equation (13.21) to Equation (13.23) correspond to Equation (13.3); (13.24)–(13.26) to (13.4); (13.27)–(13.29) to (13.5); and (13.30)–(13.32) to (13.6).

The solutions to system (13.21) to system (13.32) for $a_{i1} = 0$, and $a_{i2} = 0$ ($i = 1, \dots, 4$) will be the equilibrium points, and those with any $a_{ij} \neq 0$, with $j \neq 0$ are periodic solutions. These periodic solutions will give an approximation to the frequency and amplitude of the self-sustained oscillations of the state variables. There will be several solutions, real and imaginary, but only the real ones are of interest and thus will be retained.

The study is again divided into two cases: the case $\tau_L = 0$ and the general case $\tau_L \neq 0$.

13.5.1 Zero Load Torque Case ($\tau_L = 0$)

As stated in Section 13.3, when $\tau_L = 0$, the analysis is simpler than in the case $\tau_L \neq 0$, because in this case the system has only one equilibrium point, which is independent of the rest of the parameters.

An analytical solution of system (13.21) to system (13.32) is too complex to be given as a function of parameters ($c_1, c_2, c_4, c_5, u_2^0$), but for particular values of these parameters, a set of solutions can be given for any κ, k_p, k_i . Using the values of the parameters given in Table 13.1 the solutions to system (13.21) to system (13.32) can be obtained numerically:

- Solution 1: ($a_{10} = 1, a_{11} = 0, a_{20} = 0, a_{21} = 0, a_{22} = 0, a_{30} = 0, a_{31} = 0, a_{32} = 0, a_{40} = 0, a_{41} = 0, a_{42} = 0$) for any ω . This solution is the desired equilibrium point.
- Solution 2:

$$\begin{aligned}
 a_{10} &= \frac{4k_p + \alpha^2}{4k_p + 4\kappa k_p + k_i} & a_{11} &= 0 \\
 a_{20} &= 0 & a_{21} &= -\frac{4\beta k_p(\alpha^2 + \kappa\alpha^2 - k_i)}{k_i(4k_p + 4\kappa k_p + k_i)} & a_{22} &= -\frac{k_p\alpha\beta(4k_p + \alpha^2)}{k_i(4k_p + 4\kappa k_p + k_i)} \\
 a_{30} &= 0 & a_{31} &= 0 & a_{32} &= -\frac{\alpha\beta}{k_i} \\
 a_{40} &= 0 & a_{41} &= \beta & a_{42} &= -\frac{a_p\alpha\beta}{k_i} \\
 \omega &= \alpha
 \end{aligned}$$

where α and β are the solutions to the equations

$$k_p\alpha^4 + (4k_p^2 + 16k_p + 16\kappa k_p - 4\kappa k_i)\alpha^2 + 4k_i^2 = 0 \quad (13.33)$$

$$\begin{aligned}
 & (8\kappa k_i k_p + 32\kappa k_p^2 + 32\kappa^2 k_p^2)\beta^2 + 64\kappa^2 k_p^2 - 4\kappa k_i^2 + 4k_p^2 k_i + 4k_i + \alpha^2 k_p k_i \\
 & + 4\alpha^2 \kappa k_p^2 - 16\kappa^2 k_p k_i + 16k_p k_i + 16\kappa k_p k_i + 16\kappa k_p^3 + 64\kappa k_p^2 = 0
 \end{aligned} \quad (13.34)$$

Solution 2 is the periodic one, and the conditions that make α and β real, obtained from Equation (13.33) and Equation (13.34), are the conditions for the existence of a limit cycle.

From this general solution, it is concluded that x_1 (the quadrature rotor flux) does not oscillate, and the others state variables oscillate around its value at the equilibrium point (zero), since the bias components are null ($a_{20} = 0$, $a_{30} = 0$, and $a_{40} = 0$).

By substituting the values of κ , k_p , k_i given in Table 13.1 to verify the presence of a limit cycle, we obtain the solutions:

1. For $\kappa = 1.6$, $k_p = 0.1$, and $k_i = 1$, we obtain only the equilibrium solution
2. But for $\kappa = 1.8$, $k_p = 0.1$, and $k_i = 1$, we obtain two solutions:
 - (a) ($a_{10} = 1$, $a_{11} = 0$, $a_{20} = 0$, $a_{21} = 0$, $a_{22} = 0$, $a_{30} = 0$, $a_{31} = 0$, $a_{32} = 0$, $a_{40} = 0$, $a_{41} = 0$, a_{42}) for any ω . This solution corresponds to the equilibrium point
 - (b) ($a_{10} = 0.937$, $a_{11} = 0$, $a_{20} = 0$, $a_{21} = 0.2157$, $a_{22} = 0.0391$, $a_{30} = 0$, $a_{31} = 0$, $a_{32} = 0.4183$, $a_{40} = 0$, $a_{41} = -0.3321$, $a_{42} = 0.0418$, $\omega = 1.2595$)

If these solutions are now compared with the simulations presented in Section 13.4, it can be stated that the frequency of the limit cycle predicted earlier ($\omega = 1.2595$ rad/sec) is very similar to the one of simulation B of Figure 13.6 ($\omega = 1.3$ rad/sec). On the other hand, if the evolution of x_1 in simulation A were plotted, it could be seen that x_1 does not oscillate as is predicted in this study. The predicted amplitudes of the limit cycles are not exactly the ones obtained in the simulations because of the approximate nature of the method. These differences would be smaller if more harmonics were used, but that would make the problem computationally more involved [24]. From a qualitative point of view, the first-harmonic balance method captures the essence of the phenomenon.

13.5.2 Nonzero Load Torque Case ($\tau_L \neq 0$)

This case is more complex because there can be several equilibria, and numerical methods have been used to give a solution to the system of Equation (13.21) to Equation (13.32) for the parameter values given in Table 13.1.

1. For the values of parameters ($c_1 = 4$, $c_2 = 4$, $c_4 = 1$, $c_5 = 1$, $u_2^0 = 1$, $\kappa = 3.7$, $k_p = 0.15$, $k_i = 1$, and $\tau_L = 0.2$), only one solution can be obtained which corresponds to the equilibrium point: ($a_{10} = 0.9657$, $a_{11} = 0$, $a_{20} = -0.1455$, $a_{21} = 0$, $a_{22} = 0$, $a_{30} = 0$, $a_{31} = 0$, $a_{32} = 0$, $a_{40} = 0.0562$, $a_{41} = 0$, and $a_{42} = 0$) for any ω .

2. For the values ($c_1 = 4, c_2 = 4, c_4 = 1, c_5 = 1, u_2^0 = 1, \kappa = 3.9, k_p = 0.15, k_i = 1, \tau_L = 0.2$), there must be a limit cycle in the state variables (Section 13.3), and, indeed, there are two solutions to the system of Equation (13.21) to Equation (13.32):
 - (a) ($a_{10} = 0.9691, a_{11} = 0, a_{20} = -0.1483, a_{21} = 0, a_{22} = 0, a_{30} = 0, a_{31} = 0, a_{32} = 0, a_{40} = 0.0534, a_{41} = 0, \text{ and } a_{42} = 0$) for any ω . This solution is the equilibrium point.
 - (b) ($a_{10} = 0.8867, a_{11} = 0.1428, a_{20} = -0.1576, a_{21} = 0.2833, a_{22} = -0.0750, a_{30} = 0, a_{31} = 0.067, a_{32} = 0.1806, a_{40} = 0.0545, a_{41} = -0.0833, a_{42} = 0.0617, \text{ and } \omega = 1.9348$). This solution corresponds to the limit cycle of simulation D in Figure 13.6 (in the simulation $\omega = 1.8 \text{ rad/sec}$).

Note that in the case $\tau_L \neq 0$, all the state variables oscillate.

13.6 Conclusions

In this chapter, we have shown that self-sustained oscillations in indirect FOC for induction motors may be due to the appearance of a Hopf bifurcation. Other causes of oscillations may exist but, for these cases, the local stability of the desired equilibrium point would be preserved. Therefore, the Hopf bifurcation is the most interesting phenomenon regarding oscillations from the practical point of view. We have given some simple rules tuning the PI gains in order to prevent Hopf bifurcation and quench these oscillations. Our theoretical results were validated with some simulation evidence. Current research is under way to test our theoretical predictions in an experimental facility.

Appendix: Motor Model

The dynamic model of a current driven induction motor expressing the rotor flux and the stator currents in a reference frame rotating at rotor speed is given by (see Ref. [4] for a detailed derivation of this model and the explanation of FOC):

$$\dot{\hat{i}}_{sq} = \frac{L_m R_r}{\sigma L_r^2} \psi_{rq} - \gamma i_{sq} + \frac{1}{\sigma} u_{sq} \quad (13.35)$$

$$\dot{\hat{i}}_{sd} = \frac{L_m R_r}{\sigma L_r^2} \psi_{rd} - \gamma i_{sd} + \frac{1}{\sigma} u_{sd} \quad (13.36)$$

$$\dot{\psi}_{rq} = -\frac{R_r}{L_r}\psi_{rq} + \frac{R_r L_m}{L_r}i_{sq} \quad (13.37)$$

$$\dot{\psi}_{rd} = -\frac{R_r}{L_r}\psi_{rd} + \frac{R_r L_m}{L_r}i_{sd} \quad (13.38)$$

$$\dot{\omega} = \frac{1}{J}(-B\omega(\tau - \tau_L)) \quad (13.39)$$

$$\dot{\theta} = \omega \quad (13.40)$$

$$\tau = \frac{n_p L_m}{L_r}(\psi_{rq}i_{sd} - \psi_{rd}i_{sq}) \quad (13.41)$$

where $\sigma \triangleq L_s - L_m^2/L_r$, $\gamma \triangleq (L_m^2 R_r + L_r^2 R_s)/\sigma L_m^2$, and

L_m	mutual inductance
L_s	stator inductance
L_r	rotor inductance
R_s	stator resistance
R_r	rotor resistance
$i_s = [i_{sq}, i_{sd}]^T$	stator currents
$\psi_r = [\psi_{rq}, \psi_{rd}]^T$	rotor flux
$u_s = [u_{sq}, u_{sd}]^T$	stator voltage
θ	rotor position
ω	mechanical speed
τ	electromagnetic torque
τ_L	load torque
n_p	pole pair number
J	moment of inertia
B	friction coefficient
q	quadrature axis
d	direct axis

In current-fed motors, high-gain fast-analog current loops are used, that is

$$u_{sq} = \frac{1}{\varepsilon} \left(i_{sq}^d - i_{sq} \right) \quad (13.42)$$

$$u_{sd} = \frac{1}{\varepsilon} \left(i_{sd}^d - i_{sd} \right) \quad (13.43)$$

with (i_{sq}^d, i_{sd}^d) denoting the desired values for the stator currents, and ε a small positive number. Under these conditions, it is reasonable then to assume that the motor dynamics is described by the reduced model

obtained by setting $\varepsilon = 0$, that is

$$\dot{\psi}_r = -\frac{R_r}{L_r}\psi_r + \frac{R_r L_m}{L_r}i_s \quad (13.44)$$

$$J\dot{\omega} = -B\omega + (\tau - \tau_L) \quad (13.45)$$

$$\dot{\theta} = \omega \quad (13.46)$$

$$\tau = \frac{n_p L_m}{L_r} i_s^T \mathbf{J} \psi_r \quad (13.47)$$

where

$$\mathbf{J} \triangleq \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Now let us introduce the change of coordinates

$$x = e^{-J\rho_d} \psi_r \quad (13.48)$$

$$v = e^{-J\rho_d} i_s \quad (13.49)$$

to express the rotor flux and the stator currents in a rotating reference frame, where the rotation matrix is

$$e^{-J\rho_d} = \begin{bmatrix} \cos \rho_d & \sin \rho_d \\ -\sin \rho_d & \cos \rho_d \end{bmatrix} \quad (13.50)$$

and ρ_d can be interpreted as the angle of the desired rotor flux, and $\dot{\rho}_d$ as the slip frequency.

Applying the change of coordinate (13.48) and coordinate (13.49) into Equation (13.44) gives

$$\dot{x} = -\frac{1}{T_r}x + \frac{L_m}{T_r}v - \mathbf{J}\dot{\rho}_d x \quad (13.51)$$

where the rotor flux time constant $T_r \triangleq L_r/R_r$.

Equation (13.45) becomes

$$\dot{\omega} = \frac{1}{J}(B\omega + (\tau - \tau_L)) \quad (13.52)$$

$$\tau = \frac{n_p L_m}{L_r} v^T \mathbf{J} x \quad (13.53)$$

Defining $u_1 = \dot{\rho}_d$, $u_2 = v_1$, and $u_3 = v_2$, these equations can be written as

$$\dot{x}_1 = -c_1 x_1 + u_1 x_2 + c_2 u_2 \quad (13.54)$$

$$\dot{x}_2 = -c_1 x_2 - u_1 x_1 + c_2 u_3 \quad (13.55)$$

$$\dot{\omega} = -c_3 \omega + c_4 [c_5 (u_3 x_1 - u_2 x_2) - \tau_L] \quad (13.56)$$

where

$$c_1 = \frac{1}{T_r} \quad c_2 = \frac{L_m}{T_r} \quad c_3 = \frac{B}{J} \quad c_4 = \frac{1}{J} \quad c_5 = \frac{n_p L_m}{L_r}$$

In this chapter, c_3 is disregarded because the friction is usually very small. In speed regulation applications, the indirect FOC strategy is usually applied along with a PI speed loop. This control strategy is described by the equations

$$u_1 = \hat{c}_1 \frac{u_3}{u_2} \quad (13.57)$$

$$u_2 = u_2^0 \quad (13.58)$$

$$u_3 = k_p (\omega_{\text{ref}} - \omega) + k_i \int_0^t (\omega_{\text{ref}}(\eta) - \omega(\eta)) d\eta. \quad (13.59)$$

where $\hat{c}_1$ is an estimate for the inverse rotor time constant; c_1 , k_p , and k_i are the gains of the PI speed controller; ω_{ref} is the constant reference velocity; and u_2^0 is the constant desired value of the rotor flux norm.

We define $\kappa \triangleq \hat{c}_1 / c_1$ as a degree of tuning in the estimation of the rotor time constant. The closed-loop system (13.54)–(13.56) with the control (13.58)–(13.60) is a fourth-order system that can be expressed as

$$\dot{x}_1 = -c_1 x_1 + \frac{\kappa c_1}{u_2^0} x_2 x_4 + c_2 u_2^0 \quad (13.60)$$

$$\dot{x}_2 = -c_1 x_2 - \frac{\kappa c_1}{u_2^0} x_1 x_4 + c_2 x_4 \quad (13.61)$$

$$\dot{x}_3 = -c_4 [c_5 (x_1 x_4 - x_2 u_2^0) - \tau_L] \quad (13.62)$$

$$\dot{x}_4 = k_i x_3 - k_p c_4 [c_5 (x_1 x_4 - x_2 u_2^0) - \tau_L] \quad (13.63)$$

where the new state variables $x_3 \triangleq \omega_{\text{ref}} - \omega$ and $x_4 \triangleq u_3$.

References

1. B.K. Bose, *Power Electronics and AC Drives*, Prentice Hall, 1986.
2. S.M. Chhaya, Expert system based automated design of AC drive systems, *PhD. Dissertation*, University of Tennessee, Knoxville, 1995.
3. W. Leonhard, *Control of Electrical Drives*, Springer-Verlag, Berlin, 1985.
4. R. Ortega, A. Loria, P.J. Nicklasson, and H. Sira-Ramirez, *Passivity-based control of Euler-Lagrange systems*, Springer-Verlag, Berlin, Commun. Contr. Eng., 1998.
5. A.S. Bazanella and R. Reginatto, Robustness margins for indirect field-oriented control of induction motors, *Proceedings of the 37th CDC*, Tampa, Florida, 1998.
6. P.A.S. de Wit, R. Ortega, and I.M.Y. Mareels, Indirect field-oriented control of induction motors is robustly globally stable, *Automatica*, 32 (10), 1393–1402, 1996.
7. A.S. Bazanella, R. Reginatto, and R. Valiati, On Hopf bifurcations in indirect field oriented control of induction motors: designing a robust PI controller, *Proceedings of the 38th CDC*, Phoenix, Arizona, 1999.
8. E.H. Abed, H.O. Wang, and A. Tesi, *Control of bifurcations and chaos*, in *The Control Handbook*, W.S. Levine, Ed., IEEE Press, 1996, pp. 951–966.
9. J. Alvarez, E. Curiel, and F. Verduzco, Complex dynamics in classical control systems, *Syst. Control Lett.*, 31, 1997, pp. 277–285.
10. J. Aracil, K. Åström, and D. Pagano, Global bifurcations in the Furuta pendulum, *NOLCOS'98*, Enschede, The Netherlands, 1998, pp. 35–40.
11. J. Llibre and E. Ponce, Global first harmonic bifurcation diagram for odd piecewise linear control systems, *Dyn. Stabil. Syst.*, 11 (1), 1996, 49–88.
12. S.H. Strogatz, *Nonlinear Dynamics and Chaos*, Addison-Wesley, 1995.
13. J. Hale and H. Koçak, *Dynamics and Bifurcations*, Springer-Verlag, 1991.
14. Y.A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, 1995.
15. J. Moiola and G. Chen, *Hopf Bifurcation Analysis: A Frequency Domain Approach*, World Scientific, 1996.
16. W.M. Liu, Criterion of Hopf bifurcations without using eigenvalues, *J. Math. Anal. Appl.*, 182, 250–256, 1994.
17. M. Basso, R. Genesio, and A. Tesi, A frequency method for predicting limit cycle bifurcations, *Nonlinear Dyn.*, 13, 339–360, 1997.
18. R. Genesio and A. Tesi, Harmonic balance methods for the analysis of chaotic dynamics in nonlinear systems, *Automatica*, 28 (3), 531–548, 1992.
19. A.R. Bergen, L.O. Chua, A.I. Mees, and E.W. Szeto, Error Bounds for General Describing Function Problems, *IEEE-CAS*, 29 (6), 345–354, 1982.
20. A.I. Mees, *Dynamics of Feedback Systems*, Wiley, 1981.
21. R.G. Bartle and D.R. Sherbert, *Introduction to Real Analysis*, John Wiley and Sons, New York, 1982.
22. P.A. Cook, *Nonlinear Dynamical Systems*, 2nd ed., Prentice Hall, 1994.

23. F. Bonani and M. Gilli, Analysis of stability and bifurcations of limit cycles in Chua's circuit through the Harmonic-Balance method, *IEEE Trans. Circuits Syst.*, 46 (8), 881–890, 1990.
24. F. Salas, T. Alamo, and J. Aracil, Detecting periodic orbits in nonlinear systems, *XIX Jornadas de Automática*, in Spanish, Madrid, pp. 99–104, 1998.

14

Implementation of the Chua's Circuit and its Application in the Data Transmission

L. Boutat-Baddas, J.P. Barbot, and R. Tauleigne

CONTENTS

14.1 Introduction	504
14.2 Chua's Circuit	505
14.2.1 Realization of the Non-linear Dipole	506
14.3 Observer and Synchronization	508
14.3.1 Linearizable Error Dynamics Case	510
14.3.2 Step-by-Step Sliding Mode Observer	511
14.3.2.1 Simulation Results	512
14.3.3 Output Conditions for Observer Design	515
14.3.3.1 Simulation Results	515
14.4 Transmission by Chaotic Parameter Modulation	517
14.4.1 Simulation Results	521
14.4.2 Data Transmission Examples	521
References	524

In this chapter we highlight the efficiencies of the method proposed in chapter 9. For this, we describe the realization of an electric circuit with chaotic behavior. This circuit proposed by Chua, contains a non linear resistance decomposed in five segments. In this context, we are interested by the observers design, in order to synchronize two chaotic systems, with the aim of application to protection of data transmission.

In the first part of this chapter, we introduce the Chua's circuit and present the methodology for constructing a nonlinear resistor and discuss the experimental and simulation results of the Chua's circuit.

In the second part, we synchronize two Chua's circuits through two output variables. The first one allows a linearization of the observation error dynamics whereas the second one does not. Finally, we develop a method of chaotic transmission by parameter modulation of the Chua's circuit. Different programs and simulations are available [27].

14.1 Introduction

Since 1983, in the electronic circuits domain, Chua questions about chaos and the synchronization, which he approaches with piecewise linear electrical circuits [25]. Some years later, Pecora and Carroll clarified this approach [13, 14]. In 1990, Parlitz proposed the coupling of two identical strange attractors with a purpose to demonstrate the possibility of masking a confidential message by stacking it in a chaotic signal, and to assure its recovery after reception. This application is based on the synchronization of two chaotic systems, produced with almost identical circuits (see Figure 14.1).

In this chapter we adopt an approach based on the inclusion of information in the chaotic system structure (transmitter). Consequently, information does not evolve separately from chaos, instead it will be constantly handled by the system dynamics. This approach is called chaotic parameter modulation [23, 24, 26] (shown in Figure 14.2).

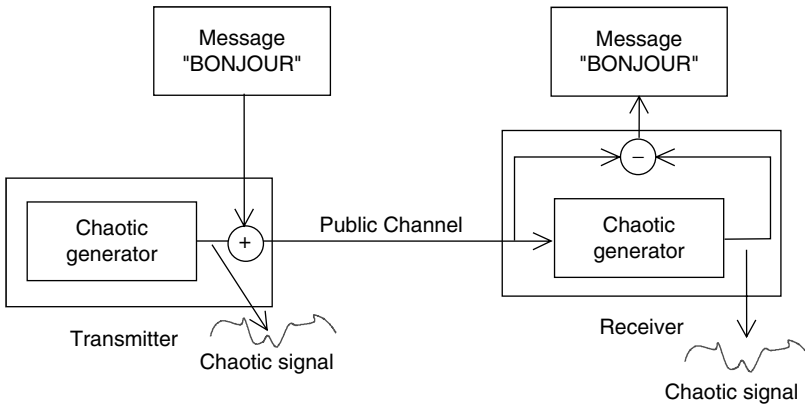


FIGURE 14.1
Additive chaos masking.

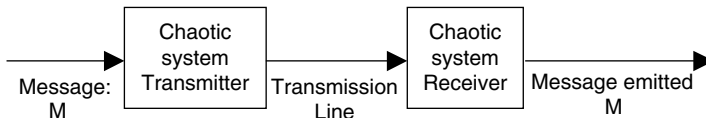


FIGURE 14.2
The chaotic parameter modulation.

The process of deciphering using the method one proposes consists in observer design with unknown input (here the message); allowing to recover the confidential message from the only information emitted by the transmitter to the receiver. However, a question arises, namely, how does one design an observer for the nonlinear systems with singularity observability? We have chosen the sliding modes observers to increase the busy band of the message and to take into account the structural changes of the system. Indeed, the loss of observability implies a change in the structure of the observer, which becomes, at the critical moment, an estimator that does not use bad information. The sliding modes observers have the advantages of converging at finite time when the system is observable, passing in a structure possessing a part without the return of information when the system is unobservable (estimator) and hanging on account of limited perturbations. This, allows us to be more close to stemming the constraints of nonlinear functions and of perturbations because it is the only type of observer which verifies the observability matching condition.

We have introduced the same concept for observability [3] (singularity observability) and we naturally recover some properties as universal inputs [6], resonant terms [21] and so on.

14.2 Chua's Circuit

We can consider the Chua's circuit as an evolution of Van der Pol's model, to which an RC circuit is added (Figure 14.3). Similarly, it contains a parallel resonator and a nonlinear resistance. Figure 14.4, gives explicitly the nonlinear current–voltage characteristic of the negative resistance.

To realize this negative resistance, Chua proposed an active circuit specially designed for this purpose [15]. For the purpose of education, we preferred to realize this dipole with the more available components (operational amplifiers) [16].

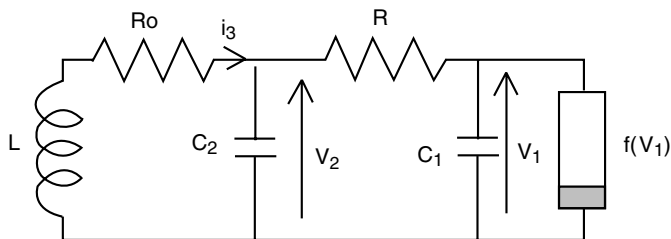
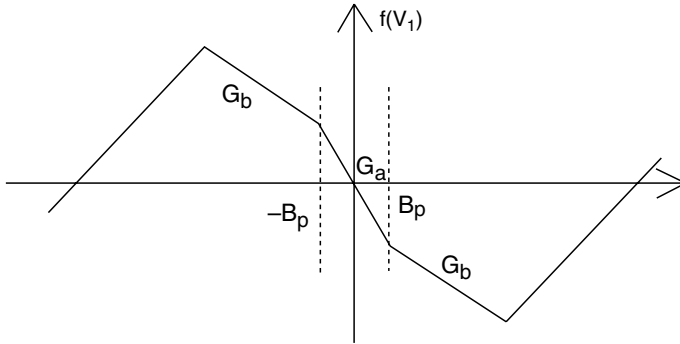


FIGURE 14.3
Chua's circuit.

**FIGURE 14.4**

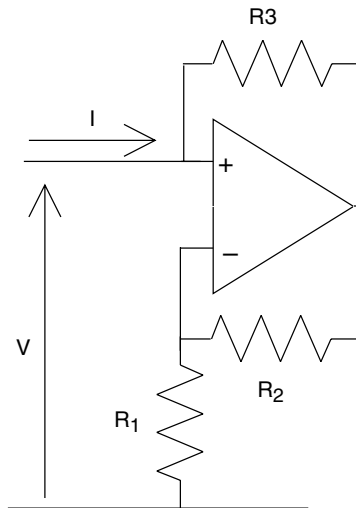
Current-voltage characteristic of the negative resistance.

14.2.1 Realization of the Non-linear Dipole

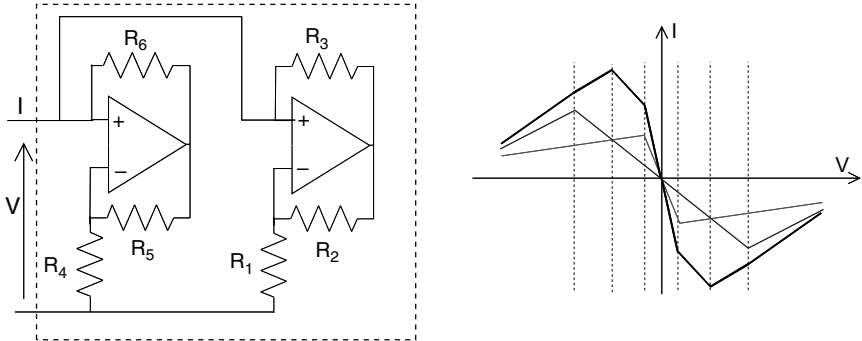
With an operational amplifier, it is simple to obtain an effect of negative resistance (see Figure 14.5). The value of the negative resistance depends on the gain of the amplifier, and the effect stops when the output voltage reaches the supply voltage.

The theoretical value of the negative resistance is:

$$R_N = \frac{I_d}{V_d} = -\frac{R_1 \cdot R_3}{R_2}$$

**FIGURE 14.5**

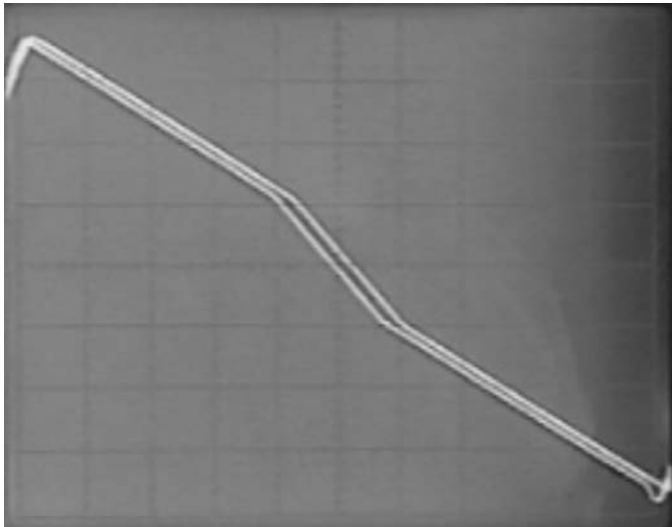
Elaboration of a negative resistance.

**FIGURE 14.6**

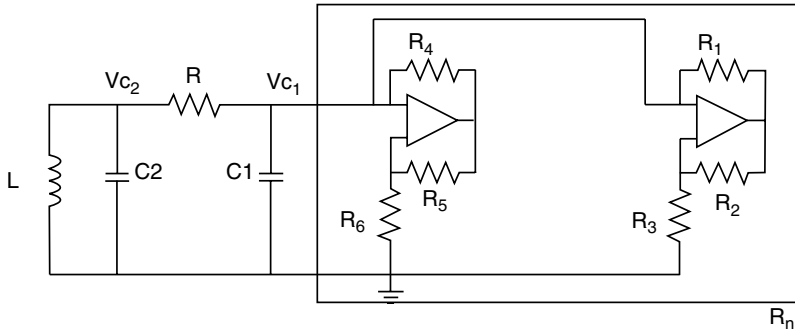
The negative resistance with double slope.

We realized two negative resistances according to this scheme connected in parallel manner (see Figure 14.6). This superposing leads to four different values for the resistances, of which two are negative, because the saturation of amplifiers does not take place at the same input voltage. The change of slope in positive parts is unimportant for our study because chaotic behavior does not investigate these parts.

In the oscilloscope, we verify the value of the obtained slopes (Figure 14.7). We also observe an effect of hysteresis inherent to the

**FIGURE 14.7**

Real current-voltage characteristic.

**FIGURE 14.8**

Complete implementation of the Chua's circuit.

mechanisms of saturation and the other imperfections of the operational amplifiers. Our final realization of Chua's circuit is shown in Figure 14.8.

The values used are: $L = 18.8 \text{ mH}$, $C_1 = 10 \text{ nF}$, $C_2 = 100 \text{ nF}$, R is a potentiometer of $5 \text{ K}\Omega$ and $R_1 = 2.2 \text{ K}\Omega$, $R_2 = R_3 = 220 \Omega$, $R_4 = 3.3 \text{ K}\Omega$, $R_5 = R_6 = 22 \text{ K}\Omega$.

It is sufficient to modify the value of a single component to investigate the various modes of this implementation (Figure 14.8). We take R , as the control parameter. With R close to $5 \text{ k}\Omega$, in the infinite limit, negative resistance cannot supply the energy to the resonator, hence no permanent oscillation is possible. For $R = 0$, we have a Van der Pol oscillator with a negative resistance, and a unique oscillation frequency. Between these two extreme values, and decreasing R , we observe the following phenomena (see Figure 14.9).

14.3 Observer and Synchronization

The problem of a nonlinear observer design with linearization of the observation error dynamics for a class of nonlinear systems, called the output injection form, has been widely investigated. Some necessary and sufficient conditions to obtain such a form are given in [8]. Using this form, it is "easy" to design an observer. Unfortunately, geometric conditions to obtain this form are very often too restrictive with respect to the system considered.

In this section, we design an observer for the well-known Chua's circuit. We choose the following dynamical variables:

- v_1 : the voltage across capacitor C_1

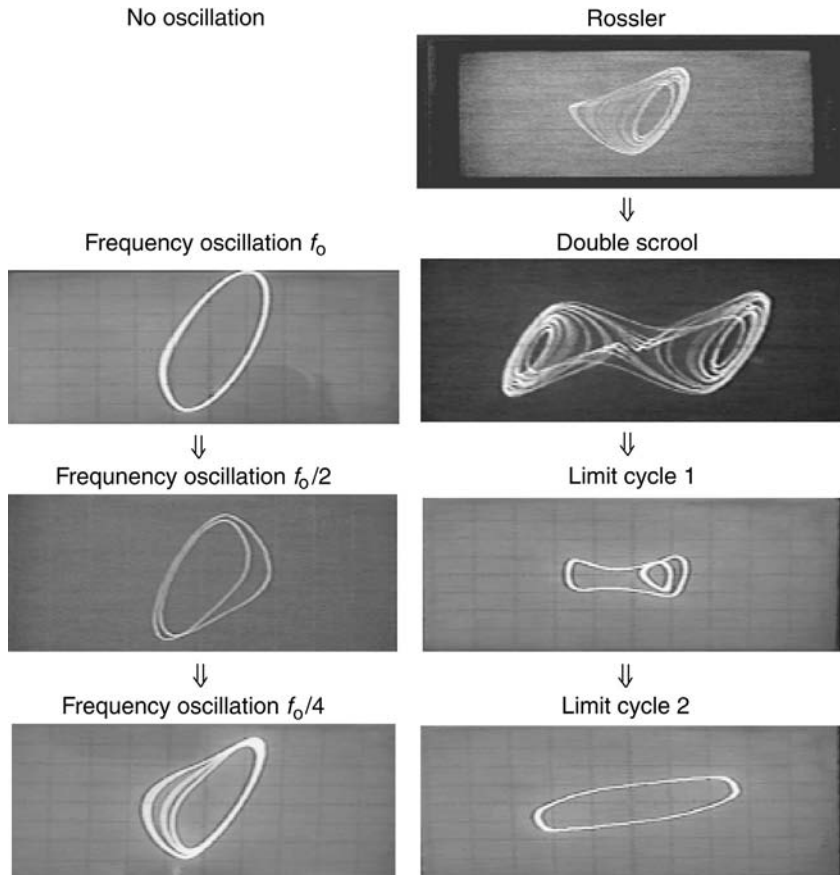


FIGURE 14.9
Route to chaos.

- v_2 : the voltage across capacitor C_2
- i_3 : the current through the inductor L

Then, according to Kirchhoff's laws:

$$\begin{aligned}
 \frac{dv_1}{dt} &= \frac{1}{C_1} \left(\frac{v_2 - v_1}{R} - f(v_1) \right) \\
 \frac{dv_2}{dt} &= \frac{1}{C_2} \left(\frac{v_1 - v_2}{R} + i_3 \right) \\
 \frac{di_3}{dt} &= \frac{1}{L} (-v_2 - R_0 i_3)
 \end{aligned} \tag{14.1}$$

with: $f(v_1) = G_b v_1 + 0.5(G_a - G_b)(|v_1 + E| - |v_1 - E|)$.

Setting $x_1 \triangleq v_1$, $x_2 \triangleq v_2$ and $x_3 \triangleq i_3$ and $x \triangleq (x_1, x_2, x_3)^T$, we obtain:

$$\begin{aligned}\frac{dx_1}{dt} &= \frac{-1}{C_1 R} (x_1 - x_2) + \frac{f(x_1)}{C_1} \\ \frac{dx_2}{dt} &= \frac{1}{C_2 R} (x_1 - x_2) + \frac{x_3}{C_2} \\ \frac{dx_3}{dt} &= \frac{-1}{L} (x_2 + R_0 x_3)\end{aligned}\tag{14.2}$$

14.3.1 Linearizable Error Dynamics Case

To observe chaotic synchronization, we realized a second Chua's circuit. First, we adjust the value of R in the first circuit, which is called the transmitter, in order to obtain the double scroll. Before coupling it to the second circuit, which is called receiver, we adjust the control of the receiver parameter (R) to also obtain the double scroll. The two circuits are then capable of working in their chaotic mode (double scroll). Of these, only one makes the coupling of two circuits (Figure 14.10). According to Parlitz's idea [4], the coupling is not bidirectional. The break of symmetry in energy exchange is due to the two operational amplifiers.

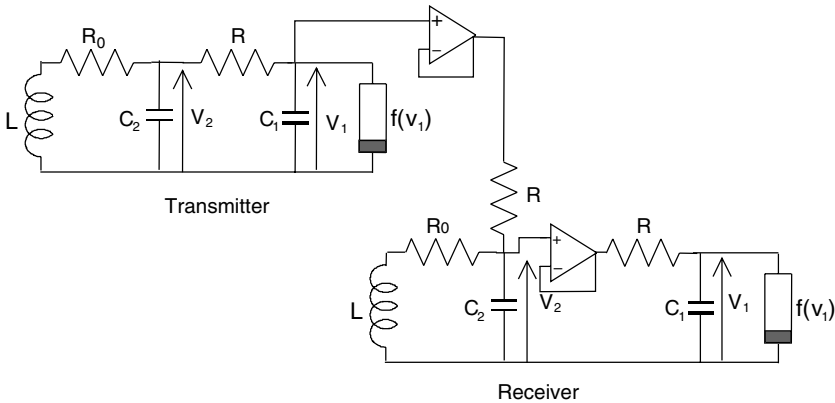


FIGURE 14.10
Parlitz's experience.

Let the output be y at state x_1 . Setting

$$A = \begin{bmatrix} \frac{-1}{C_1 R} & \frac{1}{C_1 R} & 0 \\ \frac{1}{C_2 R} & \frac{-1}{C_2 R} & \frac{1}{C_2} \\ 0 & \frac{-1}{L} & \frac{-1}{LR_0} \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

we have

$$\text{rang} \left(\begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{C_1 R} & \frac{1}{C_1 R} & 0 \\ \frac{1}{R^2 C_1^2} + \frac{1}{R^2 C_1 C_2} & -\frac{1}{R^2 C_1^2} - \frac{1}{R^2 C_1 C_2} & \frac{1}{RC_1 C_2} \end{bmatrix}$$

Then the system (14.2) is globally weakly observable [7] and linearizable by output injection. Then, there exists many observers for this system. To our knowledge, the first classical one was proposed by Parltitz [12]:

$$\begin{aligned} \frac{d\hat{x}_1}{dt} &= \frac{-1}{C_1 R} (\hat{y} - \hat{x}_2) + \frac{f(\hat{y})}{C_1} \\ \frac{d\hat{x}_2}{dt} &= \frac{1}{C_2} \left(\frac{y - \hat{x}_2}{R} + \hat{x}_3 \right) \\ \frac{d\hat{x}_3}{dt} &= \frac{1}{L} (-\hat{x}_2 - R_0 \hat{x}_3) \\ y &= x_1 \end{aligned} \tag{14.3}$$

where, $\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3)^T$ is the estimate state of x and $\hat{y}$ is the estimate output. Since then the receiver design has been changed and has become more closed to observer design [10].

14.3.2 Step-by-Step Sliding Mode Observer

Throughout this chapter we will use a step-by-step sliding mode observer [1, 17]. This kind of observer is very useful and is developed for various reasons:

- To work with reduced observation error dynamics
- For a finite time convergence of all components of the observable states

- Robustness under parameter variations is possible, if a specific condition (dual of the well-known matching condition) is verified

To do this, we provide such a kind of observer for system (14.2) with $y = x_1$ as output.

$$\begin{aligned}
 \frac{d\hat{x}_1}{dt} &= \frac{1}{C_1} \left(\frac{\hat{x}_2 - y}{R} - f(y) \right) + \lambda_1 \text{sign}(y - \hat{y}) \\
 \frac{d\hat{x}_2}{dt} &= \frac{1}{C_2} \left(\frac{y - \tilde{x}_2}{R} + \hat{x}_3 \right) + E_1 \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2) \\
 \frac{d\hat{x}_3}{dt} &= \frac{1}{L} (-\tilde{x}_2 - R_0 \tilde{x}_3) + E_2 \lambda_3 \text{sign}(\tilde{x}_3 - \hat{x}_3) \\
 y &= x_1
 \end{aligned} \tag{14.4}$$

where sign denotes the usual sign function.

With the following conditions: if $\hat{x}_1 = x_1$ then $E_1 = 1$ else $E_1 = 0$ and if $[\hat{x}_2 = \tilde{x}_2$ and $E_1 = 1]$ then $E_2 = 1$ else $E_2 = 0$. Moreover, we define the following auxiliary states:

$$\begin{aligned}
 \tilde{x}_2 &= \hat{x}_2 + E_1 C_1 R \lambda_1 \text{sign}(y - \hat{y}) \\
 \tilde{x}_3 &= \hat{x}_3 + E_2 C_2 R \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2)
 \end{aligned}$$

The proof of observation error convergence is a particular case of the proof in the last section of this paper.

REMARK 1

In practice, we add some low pass filter to the auxiliary state $\tilde{x}_i$ and we set $E_i = 1$ for $i \in \{1, 2\}$, not exactly when we are on the sliding surface but when we are close enough.

14.3.2.1 Simulation Results

Comparing the generalized phase plane of x_1, x_2 (system (14.2)) and $\hat{x}_1, \hat{x}_2$ (system (14.3) dashed line) Figure 14.11, with the generalized phase plane of x_1, x_2 (system (14.2)) and $\hat{x}_1, \hat{x}_2$ (system (14.4) dashed line) Figure 14.12, we note that the state of classical observer (system (14.3)) converges more quickly than the state of the step-by-step observer (system (14.4)). This was confirmed by Figure 14.13 and Figure 14.14, where the observation error was shown for the classical and the step-by-step observer respectively.

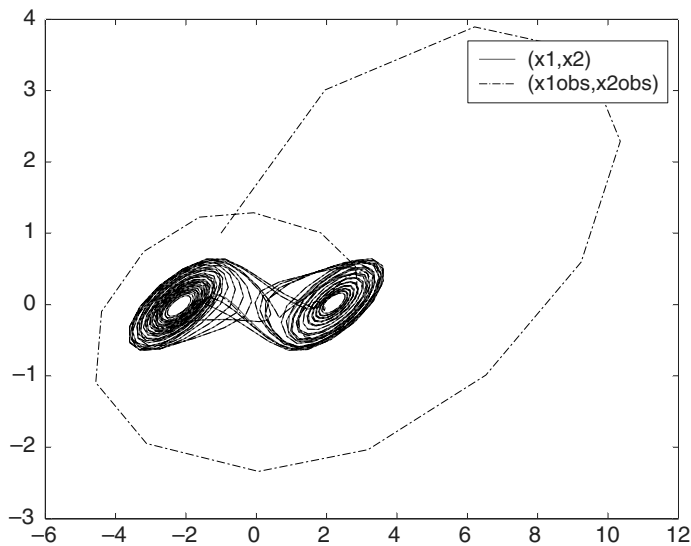


FIGURE 14.11
Double scroll attractor for system (14.2) and system (14.3).

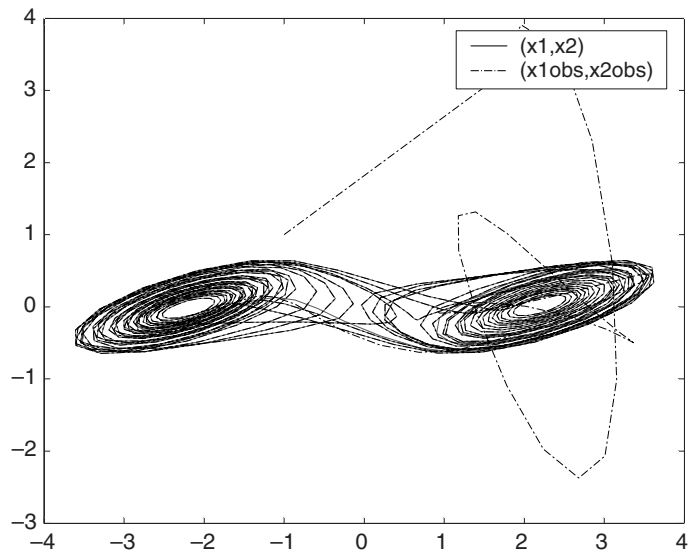


FIGURE 14.12
Double scroll attractor for system (14.2) and system (14.4).

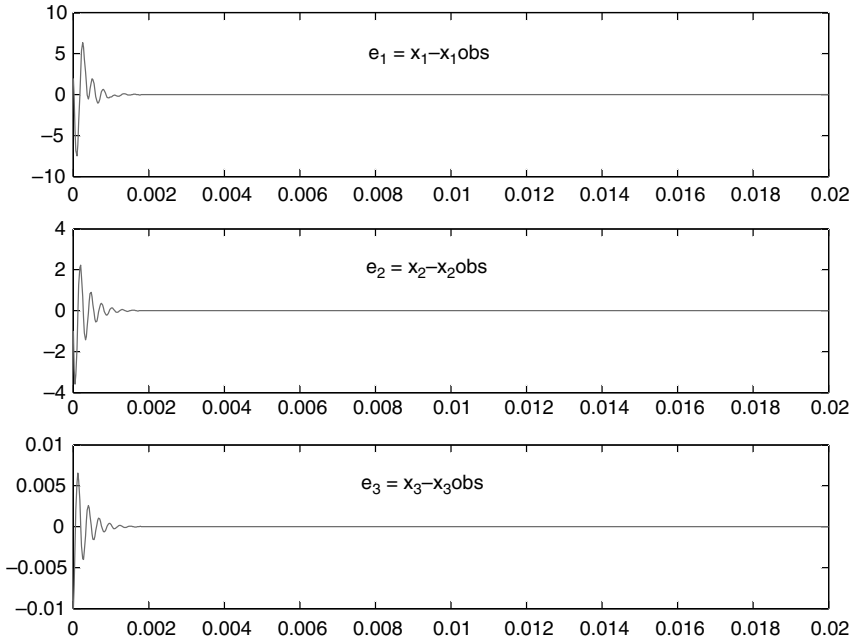


FIGURE 14.13

Observation error for system (14.2) and system (14.3).

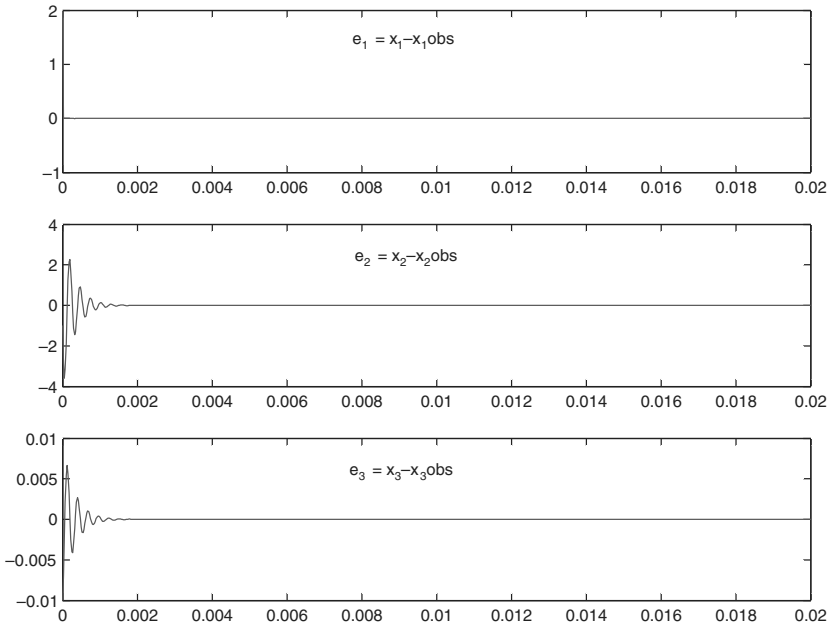


FIGURE 14.14

Observation error for system (14.2) and system (14.4).

14.3.3 Output Conditions for Observer Design

The linearization by output injection [8, 9] is a usual tool to design an observer and consequently to resolve the synchronization problem. This approach to synchronization is valid if the nonlinearity of a system depends only on the output. However, by considering another output, it is possible under some conditions to design a step-by-step sliding mode observer [1, 17, 22] in spite of the fact that linearization by output injection is not possible.

Following the introduction of the concept of a generalized output injection form [2, 18], a very interesting relationship between a chaotic system and a generalized Hamiltonian system was identified [19]. Unfortunately, considering Equation (14.2), with x_3 as output instead of x_1 , nonlinearity is not a function of the output and the result regarding output injection [2, 18, 19] cannot be used to design an observer.

Nevertheless, the observer matching condition [17] was verified (i.e., the nonlinearity $f(x_1)$ is in $\ker(C, CA)$). Therefore, it is possible to design the following step-by-step sliding mode observer:

$$\begin{aligned}
 \frac{d\hat{x}_1}{dt} &= \frac{1}{C_1} \left(\frac{\tilde{x}_2 - \hat{x}_1}{R} - f(\tilde{x}_1) \right) + E_2 \lambda_1 \text{sign}(\tilde{x}_1 - \hat{x}_1) \\
 \frac{d\hat{x}_2}{dt} &= \frac{1}{C_2} \left(\frac{\hat{x}_1 - \tilde{x}_2}{R} + x_3 \right) + E_3 \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2) \\
 \frac{d\hat{x}_3}{dt} &= \frac{1}{L} (-\hat{x}_2 - R_0 x_3) + \lambda_3 \text{sign}(x_3 - \hat{x}_3) \\
 y &= x_3
 \end{aligned} \tag{14.5}$$

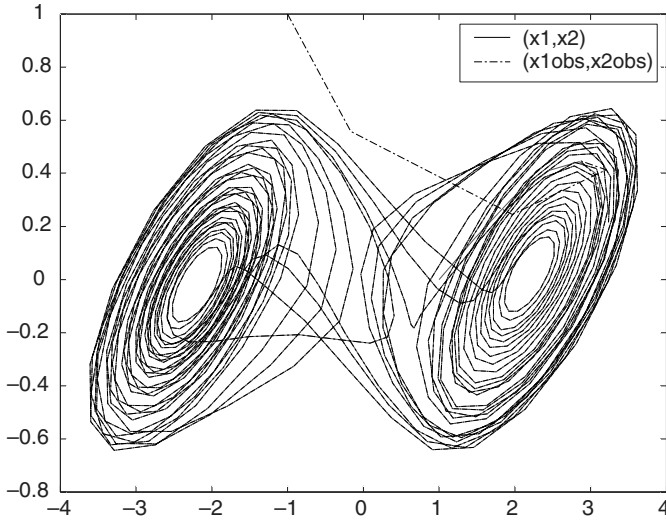
with the following conditions: if $x_3 = \hat{x}_3$ then $E_3 = 1$ else $E_3 = 0$, and if $[\tilde{x}_2 = \hat{x}_2$ and $E_3 = 1]$ then $E_2 = 1$ else $E_2 = 0$. Moreover by definition:

$$\begin{aligned}
 \tilde{x}_2 &= \hat{x}_2 - E_3 L \lambda_3 \text{sign}(x_3 - \hat{x}_3) \\
 \tilde{x}_1 &= \hat{x}_1 + E_2 C_2 R \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2)
 \end{aligned}$$

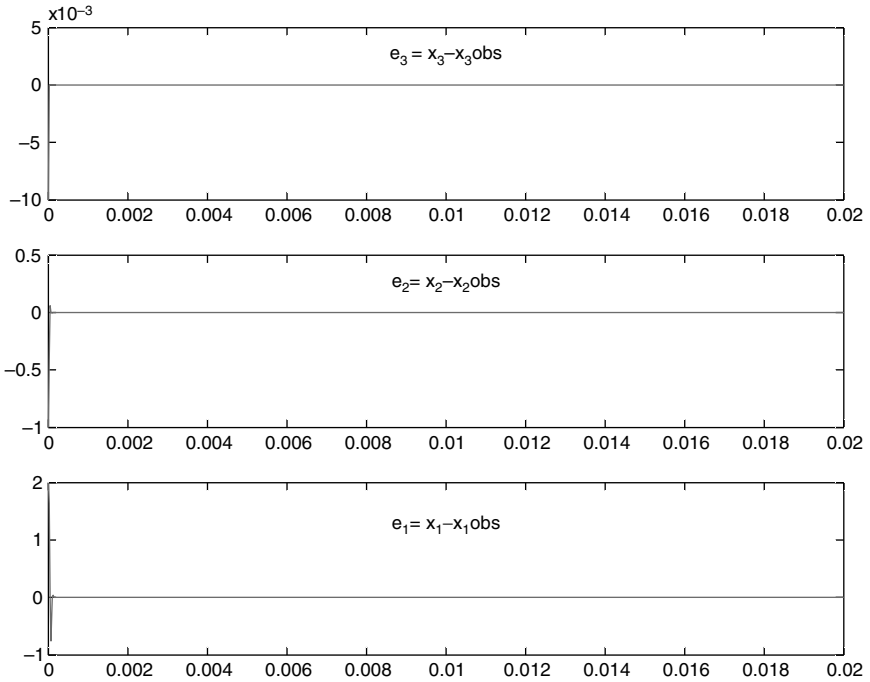
The observation error convergence was also similarly proved as in the last section.

14.3.3.1 Simulation Results

Figure 14.15 and Figure 14.16 highlight the efficiency of the step-by-step observer for the system (14.2) with x_3 as the output. It is observed that the simulation results are very close to the one previously obtained with x_1 as the output.

**FIGURE 14.15**

Double scroll attractor for system (14.2) and system (14.5).

**FIGURE 14.16**

Observation error for system (14.2) and system(14.5).

This section can be summarized in to two parts:

- Considering x_1 as output, using linearization by output injection [8], it is possible to observe the full state of the Chua's circuit.
- Thanks to a step-by-step sliding mode observer, it is also possible to design a full state observer for the Chua's circuit with a new output x_3 . For this output, the system is just observable but not linearizable by output injection.

14.4 Transmission by Chaotic Parameter Modulation

To increase the security of transmission, we propose to add some observability bifurcations at the transmission by synchronization of a chaotic system. Here, we provide an illustrative example: consider system (14.2) with x_1 as output but with $x_4 = 1/L$ as a new state. The variation of L is the information to be transmitted to the receiver. Moreover, we assume that there exists K_1 and K_2 such that $|x_4| < K_1$ and $|dx_4/dt| < K_2$, which means that the information signal and its variations are bounded. Thus, from these assumptions, we obtain the following systems:

$$\begin{aligned}
 \frac{dx_1}{dt} &= \frac{-1}{C_1 R} (x_1 - x_2) + \frac{f(x_1)}{C_1} \\
 \frac{dx_2}{dt} &= \frac{1}{C_2 R} (x_1 - x_2) + \frac{1}{C_2} x_3 \\
 \frac{dx_3}{dt} &= -(x_2 + R_0 x_3) x_4 \\
 \frac{dx_4}{dt} &= \sigma \\
 y &= x_1
 \end{aligned} \tag{14.6}$$

with σ , as an unknown bounded function (i.e., $|\sigma| < K_2$).

This system has one unobservable real mode and using the linear change of coordinate $z_1 = x_1$, $z_2 = (x_1/C_2 R) + (x_2/C_1 R)$, $z_3 = x_3/C_1 C_2 R$ and

$z_4 = x_4$ we obtain:

$$\begin{aligned}
 \frac{dz_1}{dt} &= \frac{-(C_1 + C_2)}{C_1 C_2 R} z_1 + z_2 + \frac{f(z_1)}{C_1} \\
 \frac{dz_2}{dt} &= z_3 + \frac{f(z_1)}{C_1 C_2 R} \\
 \frac{dz_3}{dt} &= \frac{z_1 z_4}{C_2^2 R} - \frac{z_2 z_4}{C_2} - R_0 z_3 z_4 \\
 \frac{dz_4}{dt} &= \sigma \\
 y &= z_1
 \end{aligned} \tag{14.7}$$

Equations (14.7) are of the normal form [3] with $\alpha = 0$ and resonant terms $h_{22} = h_{23} = 0, h_{14} = 1/C_2^2 R, h_{24} = 1/C_2$ and $h_{34} = -R_0$, but with σ as a perturbation and a nonsmooth output injection $(f(z_1)/C_1, f(z_1)/C_1 C_2 R, 0, 0)^T$. The observability singularity is given by $(z_1/C_2^2 R) - (z_2/C_2) - R_0 z_3 = 0$, and taking into account this singularity we can design an observer. Nevertheless, as system (14.6) has also a particular structure with $x_4 = z_4$ and $x_3 = C_2 C_1 R_0 z_3$, we can design an observer directly on the original state (the physical one). Obviously, the observability singularity is the same, the equation $-x_2 - R_0 x_3 = 0$ is equivalent to $(z_1/C_2^2 R) - (z_2/C_2) - R_0 z_3 = 0$.

So, we will use information contained in the terms $-x_4 x_2 - R_0 x_4 x_3$ in order to design a full state observer and recover information on x_4 contained in the equation of dx_3/dt .

For this, we use the following sliding mode observer:

$$\begin{aligned}
 \frac{d\hat{x}_1}{dt} &= \frac{1}{C_1} \left(\frac{\hat{x}_2 - y}{R} - f(y) \right) + \lambda_1 \text{sign}(y - \hat{x}_1) \\
 \frac{d\hat{x}_2}{dt} &= \frac{1}{C_2} \left(\frac{y - \tilde{x}_2}{R} + \hat{x}_3 \right) + E_1 \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2) \\
 \frac{d\hat{x}_3}{dt} &= \hat{x}_4 (-\tilde{x}_2 - R_0 \tilde{x}_3) + E_2 \lambda_3 \text{sign}(\tilde{x}_3 - \hat{x}_3) \\
 \frac{d\hat{x}_4}{dt} &= E_3 \lambda_4 \text{sign}(\tilde{x}_4 - \hat{x}_4) \\
 y &= x_1
 \end{aligned} \tag{14.8}$$

with the following conditions:

if $\hat{x}_1 = x_1$ then $E_1 = 1$ else $E_1 = 0$; similarly if $[\hat{x}_2 = \tilde{x}_2$ and $E_1 = 1]$ then $E_2 = 1$ else $E_2 = 0$; and finally if $[\hat{x}_3 = \tilde{x}_3$ and $E_2 = 1]$ then $E_3 = 1$ else $E_3 = 0$. Moreover, to take into account the observability singularity

($\tilde{x}_2 + R_0\tilde{x}_3 = 0$), we set $E_s = 1$ if $\tilde{x}_2 + R_0\tilde{x}_3 \neq 0$ else $E_s = 0$. By definition we take:

$$\begin{aligned}\tilde{x}_2 &= \hat{x}_2 + E_1 C_1 R \lambda_1 \text{sign}(y - \hat{x}_1) \\ \tilde{x}_3 &= \hat{x}_3 + E_2 C_2 \lambda_2 \text{sign}(\tilde{x}_2 - \hat{x}_2) \\ \tilde{x}_4 &= \hat{x}_4 - \frac{E_3 E_s}{(\tilde{x}_2 + R_0\tilde{x}_3 - 1 + E_s)} \lambda_3 \text{sign}(\tilde{x}_3 - \hat{x}_3)\end{aligned}$$

PROOF Here we implicitly assume that Equation (14.6) has bounded states (i.e., obvious due to energy consideration). Consequently, in the observer, we add saturation on the integrator in order to have a bounded state observer. From these two boundless considerations all λ_i may be easily chosen as constants [20].

- *First step:* Assuming that $E_1 = 0$ (if $E_1 = 1$ we directly move to the next step), the observation error dynamics ($e = x - \hat{x}$) is:

$$\begin{aligned}\frac{de_1}{dt} &= \frac{e_2}{C_1 R} - \lambda_1 \text{sign}(x_1 - \hat{x}_1) \\ \frac{de_2}{dt} &= \frac{e_2}{C_2 R} + \frac{e_3}{C_2} \\ \frac{de_3}{dt} &= [x_4(-x_2 - R_0 x_3)] - [\hat{x}_4(-\hat{x}_2 - R_0 \hat{x}_3)] \\ \frac{de_4}{dt} &= 0\end{aligned}$$

Due to the finite time convergence of the sliding mode, there exists $\tau_1 \geq 0$ such that $\forall t \geq \tau_1$, $\hat{x}_1 = x_1$ and we move to the next step.

- *Second step:* As $\hat{x}_1 = x_1$ then $E_1 = 1$ and as $e_1 = 0$ for all $t \geq \tau_1$ then $\dot{e}_2 = 0$ and consequently, invoking the equivalent vector [20], $\tilde{x}_2 = x_2$, we obtain

$$\begin{aligned}\frac{de_1}{dt} &= \frac{e_2}{C_1 R} - \lambda_1 \text{sign}(x_1 - \hat{x}_1) = 0 \\ \frac{de_2}{dt} &= \frac{e_3}{C_2} - \lambda_2 \text{sign}(x_2 - \hat{x}_2) \\ \frac{de_3}{dt} &= [x_4(-x_2 - R_0 x_3)] - [\hat{x}_4(-\hat{x}_2 - R_0 \hat{x}_3)] \\ \frac{de_4}{dt} &= 0\end{aligned}$$

Due to the finite time convergence of the sliding mode, there exists $\tau_2 \geq \tau_1 \geq 0$ such that $\forall t \geq \tau_2$, $\hat{x}_2 = \tilde{x}_2 = x_2$ and we move to the next step.

- *Third step:* As $[\hat{x}_2 = x_2 \text{ and } E_1 = 1]$ then $E_2 = 1$ and as $e_2 = 0$ for all $t \geq \tau_2$ then $\dot{e}_3 = 0$ and consequently, invoking the equivalent vector, $\tilde{x}_3 = x_3$, and we obtain

$$\begin{aligned}\frac{de_1}{dt} &= \frac{e_2}{C_1 R} - \lambda_1 \text{sign}(x_1 - \hat{x}_1) = 0 \\ \frac{de_2}{dt} &= \frac{e_3}{C_2} - \lambda_2 \text{sign}(x_2 - \hat{x}_2) = 0 \\ \frac{de_3}{dt} &= -(x_2 + R_0 x_3)e_4 - \lambda_3 \text{sign}(x_3 - \hat{x}_3) \\ \frac{de_4}{dt} &= 0\end{aligned}$$

Due to the finite time convergence of the sliding mode, there exists $\tau_3 \geq \tau_2 \geq \tau_1 \geq 0$ such that $\forall t \geq \tau_3$, $\hat{x}_3 = \tilde{x}_3 = x_3$ and we move to the next step.

- *Last step:* As $[\hat{x}_3 = x_3 \text{ and } E_3 = 3]$ then $E_3 = 1$ and we obtain:

$$\begin{aligned}\frac{de_1}{dt} &= \frac{e_2}{C_1 R} - \lambda_1 \text{sign}(x_1 - \hat{x}_1) = 0 \\ \frac{de_2}{dt} &= \frac{e_3}{C_2} - \lambda_2 \text{sign}(x_2 - \hat{x}_2) = 0 \\ \frac{de_3}{dt} &= -(x_2 + R_0 x_3)e_4 - \lambda_3 \text{sign}(x_3 - \hat{x}_3) = 0 \\ \frac{de_4}{dt} &= E_s \lambda_4 \text{sign}(\tilde{x}_4 - \hat{x}_4)\end{aligned}$$

Therefore, if $E_s = 1$, then e_4 converges to zero in finite time, else $E_s = 0$ and the e_4 dynamic is frozen (the data acquisition). Nevertheless, the singularity $(x_2 + R_0 x_3)$ is local, as the transmitter is chaotic, but enough time was not provided to the singularity to alter the data acquisition substantially.

REMARK 2

$E_s = 0$ not only when there is a singularity but also when we are close to it.

To illustrate the efficiency of the method we chose to transmit the following message:

$$L(t) = L + 0.1 \sin(100t), \quad \text{with } L = 18.8 \text{ mH.}$$

14.4.1 Simulation Results

In Figure 14.17 and Figure 14.18, if we set $E_s = 0$ on a big neighborhood of the singularity manifold $(x_2 + R_0x_3)$, we lose the information on x_4 for a long time. It was observed that the convergence of the state $\hat{x}_4$ of the observer, towards x_4 of the original system (14.6) depends on the choice of E_s . To have a good convergence it is necessary to take $E_s = 0$ on a very small neighborhood of the singularity manifold $(x_2 + R_0x_3)$, as we notice in Figure 14.19 and Figure 14.20, contrary to the first two which were realized with too large a neighborhood. We visualize the double scroll of the transmitter and the receiver which are completely synchronized.

14.4.2 Data Transmission Examples

To illustrate the efficiency of the proposed method of ciphering, we consider the secure transmission scheme (Figure 14.2) with the chaotic system (14.6) as transmitter and (14.8) as receiver.

We carried out computer-based experiments which allowed us to encrypt and decrypt a given file which can be a text, image, or sound.

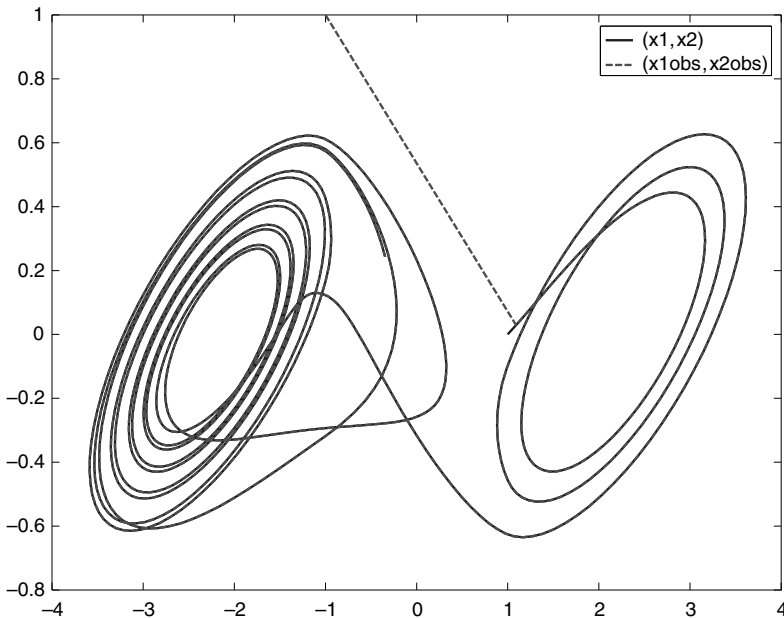
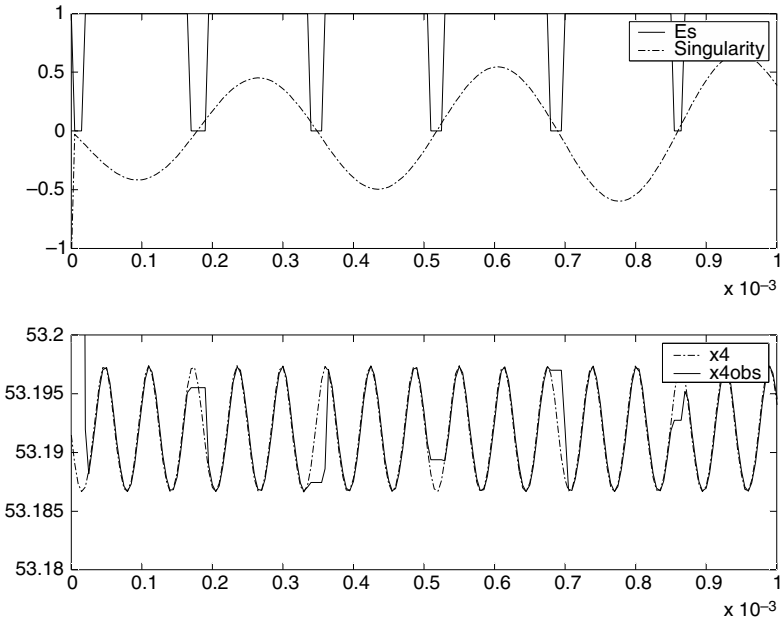
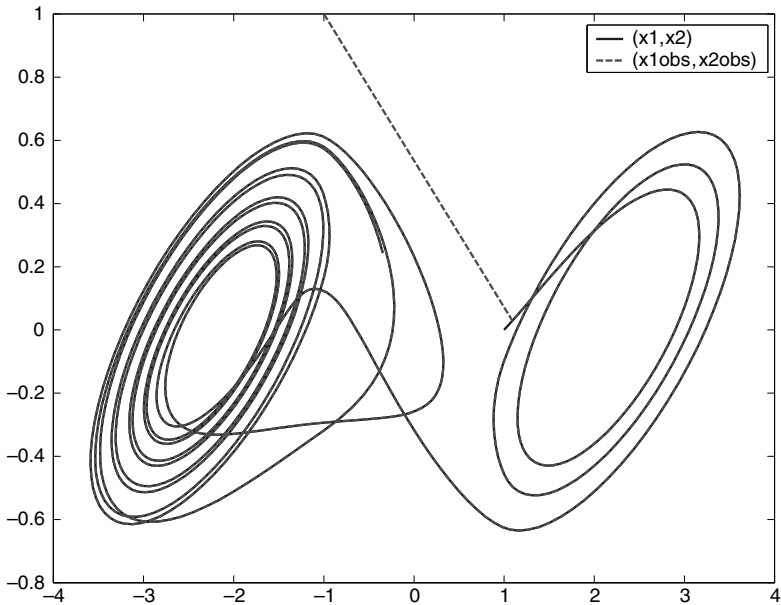


FIGURE 14.17

Double scroll attracteur for systems (14.6) and (14.8), when we set $E_s = 0$ on a big neighborhood of the singularity manifold $(x_2 + R_0x_3)$.

**FIGURE 14.18**

x_4 , $\hat{x}_4$, E_s , and the singularity $(x_2 + R_0 x_3)$, when we set $E_s = 0$ on a big neighborhood of the singularity manifold $(x_2 + R_0 x_3)$.

**FIGURE 14.19**

Double scroll attractor for systems (14.6) and (14.8), when we set $E_s = 0$ on a very small neighborhood of the singularity manifold $(x_2 + R_0 x_3)$.

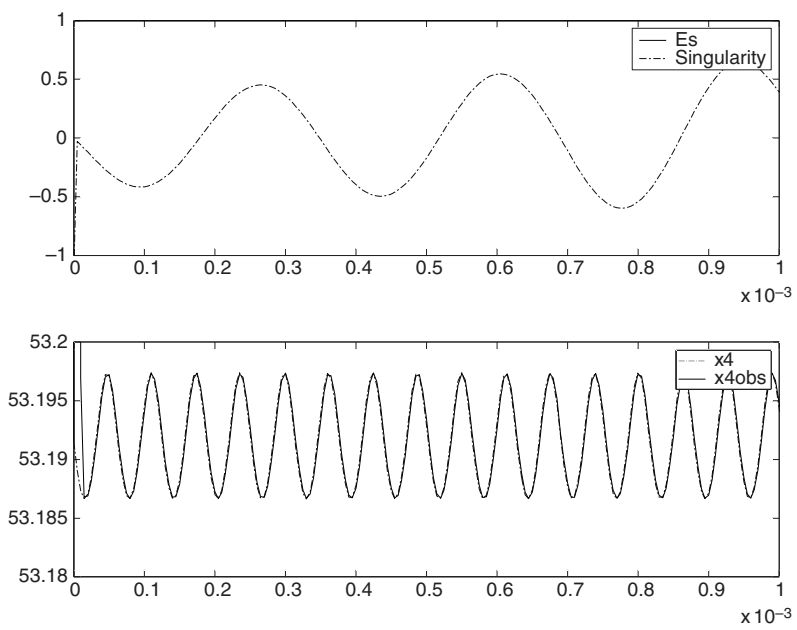


FIGURE 14.20
 x_4 , $\hat{x}_4$, E_s , and the singularity $(x_2 + R_0x_3)$, when we set $E_s = 0$ on a very small neighborhood of the singularity manifold $(x_2 + R_0x_3)$.

The encryption and decryption programs were written with the Visual C++ version 6.0. The experimental results showed that the synchronization of both systems (transmitter and receiver) was immediate and the communication result is correct and reliable Figure 14.21).

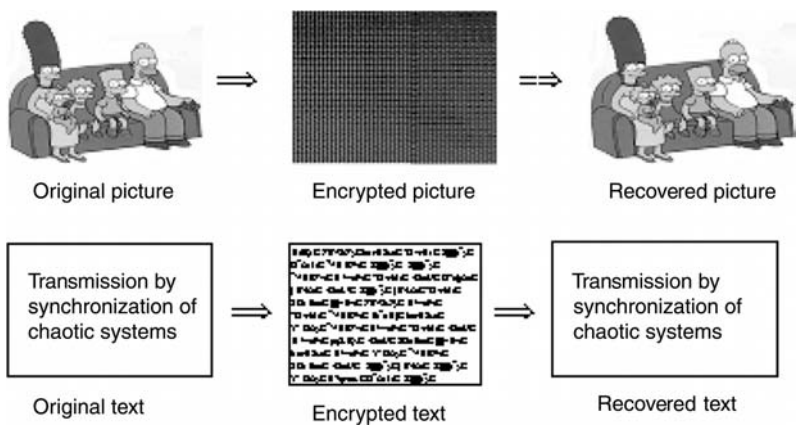


FIGURE 14.21
Transmission examples.

References

1. Barbot, J.-P., Boukhobza, T., and Djemai, M., Sliding mode observer for triangular input form, *Proceedings of the 35th CDC*, Kobe, Japan, 1996.
2. Besançon, G., A viewpoint on observability and observer design for nonlinear system, in *New directions in Nonlinear Observer Design*, Nijmeijer, H. and Fossen, T.I., *Lecture Notes in Control and Information Sciences*, 244, Springer, 1999, pp. 1–22.
3. Boutat-Baddas, L., Boutat, D., Barbot, J.-P., and Tauleigne, R., Quadratic Observability normal form, in *Proceedings of the 41st IEEE CDC* 01, 2001.
4. Cuomo, K.M., Oppenheim, A.V., and Strogatz, S.H., Synchronization of Lorenz-based chaotic circuit with application to communications, *EEE Trans. Circuit Syst.*, 40, 626–633, 1993.
5. Feldman, U., Hasler, M., and Schwartz, W., Communication by chaotic signals: the inverse system approach, in *Proceedings of the IEEE ISCAS*, Seattle, 680–683, 1995.
6. Gauthier, J.-P. and Bornard, G., Observability for any $u(t)$ of a class of bilinear systems, *IEEE TAC*, 26, 922–926, 1981.
7. Hermann, R. and Krener, A.J., Nonlinear controllability and observability, *IEEE Trans. on Autom. Contr.*, 22, 728–740, 1977.
8. Krener, A. and Isidori, A., Linearization by output injection and nonlinear observer, *Syst. Control Lett.*, 3, 47–52, 1983.
9. Krener, A. and Xiao, M.Q. Nonlinear observer design in the Siegel domain through coordinate changes, in *Proceedings of the Fifth IFAC Symposium, NOLCOS01*, Saint-Petersburg, Russia, 557–562, 2001.
10. Nijmeijer, H. and Mareels, I.M.Y., An observer looks at synchronization, *IEEE Trans. on Circuits Syst.-1: Fundam. Theory Appl.*, 44 (11), 882–891, 1997.
11. Ott, E., Grebogi, C., and Yorke, J.A., Controlling chaotic dynamical systems, in *chos: Soviet-American perspectives on nonlinear Science*, Campbell, D.K., Ed., American Institute of Physics, New York, 1990, pp. 153–172.
12. Parlitz, U., Chua, L.O., Kocarev, Lj. Halle, K.S., and Shang, A., Transmission of digital signals by chaotic synchronization, *Int. J. Bifurca. Chaos*, 2 (4), 973–997, 1992.
13. Pecora, L.M. and Carroll, T.L., Synchronizing in chaotic systems, *Phys. Rev. Lett.*, 64, 821–823, 1990.
14. Pecora, L.M. and Carroll, T.L., Synchronizing chaotic circuits, *IEEE Trans. Circuit Syst.*, 38, 453–456, 1991.
15. Cruz, J.M. and Chua, L.O., A CMOS IC nonlinear resistor for Chua's circuit, *IEEE Trans. Circuits Syst.*, 39 (12), 1992.
16. Kennedy, M.P., Robust OP amp realization of Chua's circuit, *Frequenz*, 46, 66–80, 1992.
17. Perruquetti, W. and Barbot, J.-P., *Sliding Mode control in Engineering*, Marcel Dekker, New York Basel, 2002.
18. Plestan, F. and Glumineau, A., Linearization by generalized input output injection, *Syst. Contr. Lett.*, 31, 115–128, 1997.

19. Sira-Ramirez, H. and Cruz-Hernandez, C., Synchronization of chaotic systems: a generalized Hamiltonian systems approach, *Int. J. Bifurcat. Chaos*, 11 (5), 1381–1395, 2001.
20. Utkin, V.I., *Sliding Modes in Control Optimization*, Engineering Series, Springer-Verlag, Berlin-Heidelberg, New York, 1992.
21. Wiggins, S., *Introduction to applied nonlinear dynamical systems and chaos*, in *Texts in Applied Mathematics* vol. 2, Springer, New York, 1990.
22. Xiong, Y. and Saif, M., Sliding mode observer for nonlinear uncertain systems, *IEEE Trans. Autom. Contr.*, 46, 2012–2017, 2001.
23. Yang, T. and Chua, L., Secure communication via chaotic parameter modulation, *IEEE Trans. Circuit Syst.-I: Fundam. Theory Appl.*, 43 (9), 817–819, 1996.
24. Yang, T., Wah Wu, C., and Chua, L., Cryptography based on chaotic systems, *IEEE Trans. Circuit Syst.-I: Fundamental Theory Appl.*, 44 (5), 469–472, 1997.
25. Tang, Y.S., Mees, A.I., and Chua, L.O., Synchronisation and Chaos, *IEEE Trans. Circuits Syst.*, 30 (9), 1983.
26. Zhou and Ling, X.T., Problems with the chaotic inverse system encryption approach, *IEEE Trans. Circuit Syst.-I: Fundam. Theory Appl.*, 44 (3), 268–271, 1997.
27. www.ec-lille.fr/lisac

15

Synchronization of Discrete-Time Chaotic Systems for Secured Data Transmission

I. Belmouhoub and M. Djemai

CONTENTS

15.1	Introduction	528
15.2	An Example of Discrete-Time Hyperchaotic Systems	528
15.2.1	Lyapounov Exponents	530
15.2.2	Observability Analysis	531
15.2.2.1	Linear observability	531
15.2.2.2	Quadratic observability	531
15.2.3	Observer Design	532
15.2.3.1	Computation of $\tilde{x}_2$	532
15.2.3.2	Computation of $\hat{x}_2^+$	532
15.2.4	Simulations	533
15.2.5	Discussions	533
15.3	Description of the Transmission Scheme for the DCCIM	534
15.3.1	The Transmitter	534
15.3.2	The Receiver	535
15.3.3	DOMC for Triangular Systems	535
15.3.4	The Deciphering Process	538
15.3.4.1	Step-by-Step Delayed Reconstructor Design	539
15.3.4.2	The Deciphering Algorithm	539
15.3.5	Observability Bifurcations	540
15.4	The Mandelbrot Map for the DCCIM	541
15.4.1	Structural Analysis of the Mandelbrot Map	542
15.4.1.1	Bifurcations Diagram	542
15.4.1.2	Lyapounov Exponents	543
15.4.1.3	Arnold's Tongue	545
15.4.2	Message Deciphering	546
15.4.2.1	The DOMC for the Mandelbrot Application	546
15.4.2.2	The Step-by-Step Delayed Reconstructor	546
15.4.3	Bifurcation Extending Function	547

15.5 A Ciphering–Deciphering Software	547
15.6 Conclusions	550
References	550

15.1 Introduction

The use of self-synchronizing chaotic systems for secure communications and data transmission has been an active area of research for the past few years. Although its strength for use in cryptographic type security is under debate, it has been shown that communications privacy can be enhanced by masking signals with chaotic carriers. Given the recent emergence of the so-called “keyless” crypto systems, we conclude that the popularity of chaotic communications is on the rise.

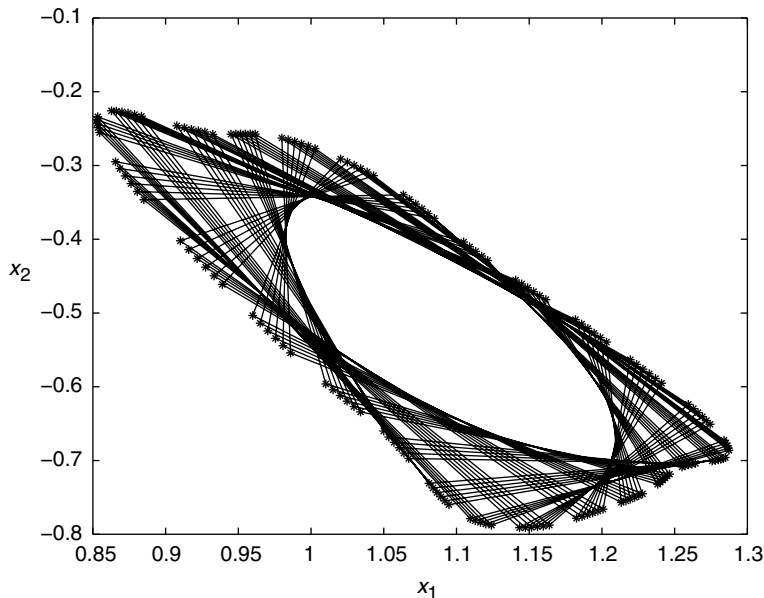
These methods are directly threatened by the expansion of computer power and parallel computation development. Consequently, new techniques appeared such as cryptography by chaos. In previous chapters, the exploitation of complex and attractive properties of chaos in promoting the characteristics of communication systems [2, 4, 7, 11, 12] was discussed. In addition, the efficiency of the so-called inclusion method (Chapter 9) in secured transmissions was approached.

In this chapter, we illustrate the *discrete-time cryptography by chaos based on the inclusion method* (DCCIM), with the background of observability normal forms and observability bifurcation analysis. The purpose is to give examples of encoding method based on the use of chaotic systems and the *Inclusion method*.

First, we present an example of chaotic synchronization, which is a capital phenomenon in the realization of a cryptographic application using chaotic processes. Then, we present the description of the transmission scheme for the DCCIM. The illustration of the method will be presented on the basis of the Mandelbrot map. Finally, some ciphering examples and conclusion ends the chapter.

15.2 An Example of Discrete-Time Hyperchaotic Systems Synchronization

In this section, we resume the Burgers map [8] (see Figure 15.1) introduced in Chapter 9. We extend the structural features of this map, to illustrate the

**FIGURE 15.1**

Phase portrait of Burgers map.

key-role of the observability normal forms in the synchronization of two hyperchaotic systems. We give, among others, the (numerical) proof of the hyperchaotic nature of this map.

REMARK 1

In Figures 15.1 and 15.2, the stars illustrate the discrete behavior of the Burgers map trajectory and the observer trajectory, respectively (each star corresponds to an iteration). However, the straight lines designate the evolution of the orbit during the iterations.

We recall the Burgers map equations:

$$\begin{cases} x_1^+ = f_1(x, p) = (1 + a) x_1 + x_1 x_2 \\ x_2^+ = f_2(x, p) = (1 - b) x_2 - x_1^2 \end{cases} \quad (15.1)$$

with $p = [a, b]^T$ and $f(x, p) := [f_1(x, p), f_2(x, p)]^T$. The output of this map is $y = x_1$.

REMARK 2

All the simulations were performed for the following parameters values $a = 0.548$ and $b = 2.28$.

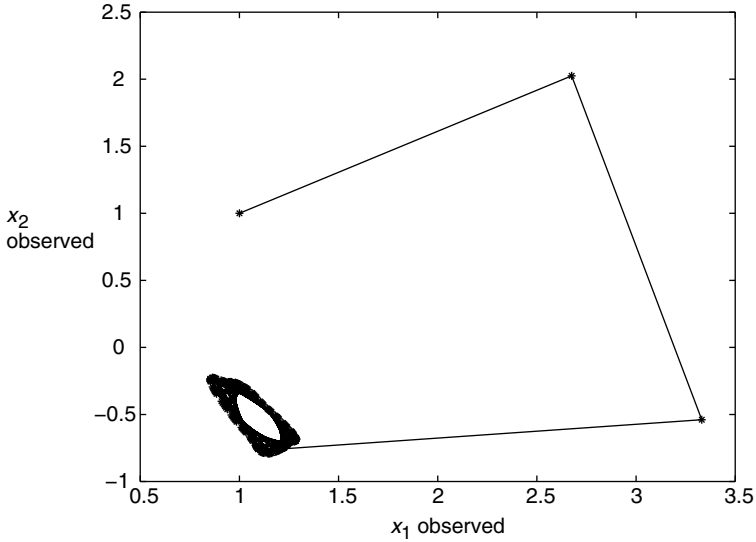


FIGURE 15.2
Phase portrait of the observer.

15.2.1 Lyapounov Exponents

We recall that the Lyapounov exponents characterize the behavior of a dynamical system. They measure the speed of exponential divergence of two nearby trajectories which stem from slightly different initial conditions. Therefore, negative Lyapounov exponents indicate stability and in the other case, a chaotic evolution, if the trajectory evolve in a bounded space.

The Burgers map possesses three equilibrium points:

$$x_{e1} = (0, 0), \quad x_{e2} = (\sqrt{ab}, -a) \quad \text{and} \quad x_{e3} = (-\sqrt{ab}, -a)$$

A Matlab simulation was implemented to compute the Lyapounov exponents $(\lambda_i(x_{e_j}))_{1 \leq i \leq 2, 1 \leq j \leq 3}$ (introduced in Chapter 2 in this book) for the three stationary points $(x_{e_j})_{1 \leq j \leq 3}$ of the Burgers map according to the following formula [3, 8]: $\forall 1 \leq i \leq 2$ for $1 \leq j \leq 3$,

$$\lambda_i(x_{e_j}) = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \log |q_i(f_o^N(x_{e_j}, p))| \right) \quad (15.2)$$

where $(q_i)_{1 \leq i \leq 2}$ are the eigenvalues of $D(f_o^k(x_{e_j}, p))$, the Jacobian matrix¹ of f_o^k evaluated in the equilibrium point x_{e_j} .

For more convenience, we use the following notation:

$$\forall 1 \leq j \leq 3, \quad \lambda(x_{e_j}) = (\lambda_1(x_{e_j}), \lambda_2(x_{e_j}))$$

Then, we obtain the following exponents (for $N = 10^5$):

$$\lambda(x_{e_1}) \simeq (0.436968, 0.436968)$$

$$\lambda(x_{e_2}) = \lambda(x_{e_3}) \simeq (0.0989662, 0.0989662)$$

It ensues from positive values of $(\lambda_i(x_{e_j}))_{1 \leq i \leq 2, 1 \leq j \leq 3}$ that the Burgers map is hyperchaotic (because unstable) in both directions $(1, 0)^T$ and $(0, 1)^T$.

REMARK 3

The Wolf, Swinney, and Vastano algorithm [3] may be used, for the computation of Lyapounov exponents. Nevertheless, owing to pedagogical concerns, we prefer to present an easy method for the implementation of (15.2).

15.2.2 Observability Analysis

15.2.2.1 Linear observability

The burgers map is linearly observable except in the following straight line:

$$L = \{(x_1, x_2) \in \mathfrak{N}^2 / x_1 = 0\}$$

15.2.2.2 Quadratic observability

The iteration of the output y , gives: $y^+ = (1 + a)x_1 + x_1x_2$. Therefore, by derivation, we obtain, the observation matrix² $J(x)$ of y and y^+ , and its determinant is:

$$|J(x)| = \begin{vmatrix} 1 & 0 \\ 1 + a + x_2 & x_1 \end{vmatrix} = x_1$$

¹The Jacobian matrix of the Burgers map is given by:

$$D(f(x, p)) = \begin{bmatrix} 1 + a + x_2 & x_1 \\ -2x_1 & 1 - b \end{bmatrix}$$

² $J = [dh^T d(f \circ h)^T \dots d(f_o^{n-1} \circ h)^T]^T$.

It ensures the quadratic observability of the Burgers map when $x_1 \neq 0$. In the opposite case, we have a local loss of observability in the x_2 direction.

We conclude that (15.1) is quadratically unobservable in one direction, in L .

In the following section, we will show how we can construct a relevant observer, thanks to this observability quadratic normal form and specifically, to the resonant term, $x_1 x_2$.

REMARK 4

For the notations, the reader should refer to Remark 8, given in Section 9.6 of Chapter 9.

15.2.3 Observer Design

We propose an observer for system (15.1), allowing to recover the states of the chaotic system, from the measured state x_1 . Thus, using delayed corrections on the second observer dynamic, it is possible to ensure the synchronization of the observer with the Burgers map. The first observer dynamic is represented by the following equation:

$$\hat{x}_1^+ = (1 + a) x_1 + x_1 \hat{x}_2 \quad (15.3)$$

15.2.3.1 Computation of $\tilde{x}_2$

Thanks to the resonant term $x_1 x_2$, the state $\tilde{x}_2$ is computed. Thus, from the equation (15.3) and the observation error dynamic $e_1^+ = x_1^+ - \hat{x}_1^+ = y e_2$. So,

$$e_2^- = \frac{e_1}{y^-}, \quad \forall y \neq 0$$

Then, throughout the bifurcation straight line L , we cannot compute $\tilde{x}_2$. This singularity is by-passed as follows:

$$\tilde{x}_2^- = \begin{cases} \hat{x}_2^- + \frac{e_1}{y^-} & \text{if } y^- \neq 0 \\ (\hat{x}_2)^{2-\varepsilon} & \text{else} \end{cases}$$

where ε is a constant in the neighborhood of zero.

15.2.3.2 Computation of $\hat{x}_2^+$

The reconstructed state $\tilde{x}_2$ (i.e., $\tilde{x}_2 = (1 - b)\tilde{x}_2^- - (x_1^-)^2$ from the second equation of (15.1)) is implemented in the observer in order to synchronize this last one with the studied chaotic system, so we compute $\hat{x}_2^+$, the

prediction of $\tilde{x}_2$ (i.e., x_2), as follows:

$$\hat{x}_2^+ = (1 - b)\tilde{x}_2 - x_1^2 \quad (15.4)$$

15.2.4 Simulations

For the simulations, the initial conditions of (15.1) and the observer (15.3), (15.4) are $X_0 = [1.05, -0.66]^T$ and $\hat{X}_0 = [1, 1]^T$, respectively, where $\hat{x}_2(0) = \tilde{x}_2(0) = 1$. The initial conditions are such that $y(0) \neq 0$. So, after one step $\tilde{x}_2^+ = x_2^+$ then, $(\hat{x}_2)^{2+} = (x_2)^{2+}$. Consequently from (15.3), after three steps, we obtain $(\hat{x}_1)^{3+} = (x_1)^{3+}$.

15.2.5 Discussions

The observer (Figure 15.2) synchronizes perfectly with the original system (Figure 15.1) after three iterations. This is confirmed by the finite-time convergence of the observation errors e_1 and e_2 . In fact, the error dynamic $e_1 = x_1 - \hat{x}_1$, converges in three iterations, whereas $e_2 = x_2 - \hat{x}_2$, converges in two steps (Figure 15.3).

In this example, we focused our attention on the synchronization of two chaotic systems. However, this observer design may be used in more general context as observer-based control. For this purpose, controllable and observable bifurcations may not be studied separately and a predictor will be requested.

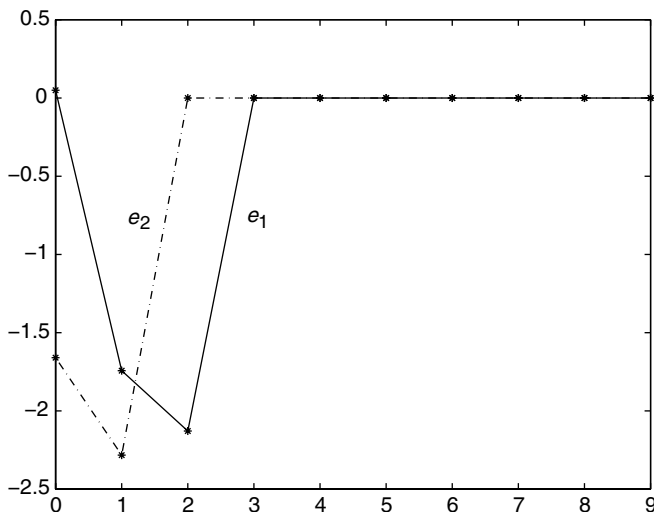


FIGURE 15.3

Observation error dynamics on x_1 and x_2 . A zoom on the first 10 iterations (10 stars).

The previous example underlines our reflection that the normal form allows to simplify the structural analysis and generally improves the observer design. This is due to the fact that the normal form has the same observability properties as each element of the equivalent class.

Now, let us describe the formalism of the DCCIM technique [1] founded on the synchronization principle.

15.3 Description of the Transmission Scheme for the DCCIM

The transmission scheme considered is a symmetric cryptosystem with private key. The key consists a part or the totality of the chaotic system parameters. This allows to exploit another aspect of the sensibility of such a system.

15.3.1 The Transmitter

It is represented by a triangular chaotic system in discrete-time, of the form:

$$\begin{aligned} x^+ &= f(x, p) + g(p)m \\ y &= x_1 \end{aligned} \tag{15.5}$$

where the state vector $x \in \mathbb{R}^n$, the parameters vector (the key) $p \in \mathbb{R}^m$. The variable y is the output of the transmitter.

The chaotic generator, $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ and the vector field $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$ are such that,

$$\begin{aligned} f(x, p) &= f_1[(x_1, x_2, p), \dots, f_i(x_1, \dots, x_{i+1}, p), \dots, f_n(x_1, \dots, x_n, p)]^T \\ g(p) &= [0, \dots, 0, g_n(p)]^T \end{aligned}$$

The variable $m \in \mathbb{R}$ is the confidential message included in the last dynamic of the ciphering process.

Hence, the transmitter is considered as a dynamical system with unknown input m . Consequently, the message deciphering consists in the resolution of the so-called left invariability problem³ by the construction of a discrete-time observer. We will prove later that the DOMC, defined

³In a usual control scheme, m is on the left side and y is on the right side of the bloc diagram. The objective is to recover the unknown input m of the system (15.5) thanks to output y , thus we employ the expression: left invertibility.

in Chapter 9 of this book, is a necessary and sufficient condition to solve this problem. In fact, it ensures the existence of a unique solution, in a neighborhood of an equilibrium point x_e of the system.

REMARK 5

For the sake of simplicity, we consider chaotic systems under triangular form. However, the DCCIM can be generalized to any system verifying the DOMC.

On the transmission line, the only information transmitted to the receiver is the data y .

15.3.2 The Receiver

It is represented by a so-called step-by-step delayed reconstructor. Its task is to recover the confidential message m .

REMARK 6

The term “delayed reconstructor” is used because the confidential message is extracted from the observation errors with a delay depending on the dimension n of the chaotic system in which it is implemented.

15.3.3 DOMC for Triangular Systems

For the triangular system (15.5), the DOMC is given by the following criteria.

PROPOSITION 1

1. $\forall i \in \{1, \dots, n-1\}$, $f_i(x, p)$ satisfies almost everywhere in $V(x_e)$ the condition:

$$\frac{\partial f_i(x, p)}{\partial x_{i+1}} \neq 0 \quad (15.6)$$

where $V(x_e) \subset \mathbb{R}^n$ is a neighborhood of x_e .

2. The term $g_n(p) \neq 0$.

PROOF Because of the triangular structure of system (15.5), $J(x, p)$ is also triangular. Hence, its determinant for $x \in V(x_e)$, can be described as follows:

$$\det(J(x, p)) = \prod_{i=1}^{n-1} \gamma_i(x, p)$$

where

$$\Upsilon_i(x, p) = \Pi_{j=0}^{n-i-1} \left(\frac{\partial f_i}{\partial x_{i+1}^{(j)+}} (x_1^{(j)+}, \dots, x_{i+1}^{(j)+}, p) \right)$$

It ensues that $\det(J(x, p)) \neq 0$, if and only if $(\partial f_i / \partial x_{i+1})(x, p) \neq 0$. Therefore, under this condition and by setting,

$$S_i = \left\{ x \in V(x_e) \text{ such that } \frac{\partial f_i}{\partial x_{i+1}}(x, p) = 0 \right\}$$

the bifurcations manifold of (15.5), J is of full rank everywhere in $V(x_e) / (\bigcup_{i=1}^n S_i)$. This corresponds to the first criterion of the DOMC.

We have

$$(J^T \cdot g)(x, p) = (0, \dots, 0, \Gamma(x, p) \cdot g_n(p))^T$$

where

$$\Gamma(x, p) = \Pi_{i=0}^{n-1} \left(\frac{\partial f_i}{\partial x_{i+1}^{(n-i-1)+}} (x_1^{(n-i-1)+}, \dots, x_{i+1}^{(n-i-1)+}, p) \right)$$

So, the second DOMC criterion is verified if the term

$$\Gamma(x, p) \cdot g_n(p) \neq 0 \quad (15.7)$$

almost everywhere around x_e .

As (15.6) is satisfied almost everywhere in $V(x_e)$, (15.7) holds if and only if $g_n(p) \neq 0$. ■

Example 1

Let us consider the triangular system:⁴

$$\begin{cases} x_1^+ = x_1 + 2x_2^2 \\ x_2^+ = x_3 + x_1x_3 \\ x_3^+ = x_1^2 - x_3^2 + 0.5m \\ y = x_1 \end{cases} \quad (15.8)$$

⁴We assume that $x_3 \neq 0$.

System (15.8) is linearly unobservable, it satisfies the DOMC if the following conditions are fulfilled:

- (i) $J = [(dy)^T, (dy^+)^T, (dy^{++})^T]^T$ is of rank 3 almost everywhere in the neighborhood of the origin.
- (ii) $Jg = (0, 0, \star)^T$ in the neighborhood of the origin is a non-null function of x almost everywhere around the origin.

Let us prove the first point. We have,

$$\begin{aligned}
 dy &= dx_1 = [1, 0, 0]^T \\
 dy^+ &= df_1(x_1, x_2, p) \\
 &= \left[\frac{\partial f_1}{\partial x_1}, \frac{\partial f_1}{\partial x_2}, \frac{\partial f_1}{\partial x_3} \right]^T = [1, 4x_2, 0]^T \\
 dy^{++} &= df_1(x_1^+, x_2^+, p) \\
 &= \left[\left(\frac{\partial f_1}{\partial x_1^+} \frac{\partial f_1}{\partial x_i} + \frac{\partial f_1}{\partial x_2^+} \frac{\partial f_2}{\partial x_i} \right)_{1 \leq i \leq 3} \right]^T \\
 &= [4x_3^2(1 + x_1), 4x_2, 4x_3(1 + x_1)^2]^T
 \end{aligned}$$

Because $\partial f_1/\partial x_3 = 0$ then,

$$|J(x)| = \left(\frac{\partial f_1}{\partial x_2} \cdot \frac{\partial f_1}{\partial x_2^+} \cdot \frac{\partial f_2}{\partial x_3} \right)(x) = 16x_2x_3(1 + x_1)^2$$

Hence, J is of rank 3 if and only if $\partial f_1/\partial x_2 \neq 0$ and $\partial f_2/\partial x_3 \neq 0$ (i.e., $x_2 \neq 0$, resp. $1 + x_1 \neq 0$). This illustrates the first result of Proposition 1.

Let us define the bifurcations manifolds:

$$\begin{aligned}
 S_1 &= \left\{ x \in V(0) \subset \mathbb{R}^3 \quad \text{such that} \quad \frac{\partial f_1}{\partial x_2}(x_1, x_2) = 4x_2 = 0 \right\} \\
 S_2 &= \left\{ x \in V(0) \subset \mathbb{R}^3 \quad \text{such that} \quad \frac{\partial f_2}{\partial x_3}(x_1, x_2, x_3) = 1 + x_1 = 0 \right\}
 \end{aligned}$$

Then, J is of full rank everywhere in $V(0) \setminus (S_1 \cup S_2)$.

As for the second point, we have

$$J \cdot g = [0, 0, (J \cdot g)_3]^T$$

where

$$\begin{aligned}
 (J \cdot g)_3 &= g_3 \cdot (\partial f_1/\partial x_2^+)(x_1^+, x_2^+) \cdot (\partial f_2/\partial x_3)(x_1, x_2, x_3) \\
 &= 4x_3(1 + x_1)^2
 \end{aligned}$$

If $g_3 = 0.5 \neq 0$ then, $(J \cdot g)_3 \neq 0$ everywhere in $V(0) \setminus (S_1 \cup S_2)$.

Now, the nonlinear observability is formulated as follows:

DEFINITION 1

- If $\forall i \in \{1, \dots, n\}$, $S_i = \emptyset$; system (15.5) is said to be locally nonlinearly observable.
- If $\forall i \in \{1, \dots, n\}$, S_i is of null measure and if $\exists i \in I \subset \{1, \dots, n\}$, such that $S_i \neq \emptyset$; system (15.5) is said to be locally almost everywhere nonlinearly observable.

REMARK 7

According to this definition system (15.5) is locally almost everywhere (i.e., in $V(0) \setminus (S_1 \cup S_2)$) nonlinearly observable.

From the Proposition 1, ensues the following result:

COROLLARY 1

If system (15.5) is (almost) everywhere nonlinearly observable; the information m is recovered (almost) everywhere provided $g_n \neq 0$.

PROPOSITION 2

The left invariability problem admits a unique solution m almost everywhere around x_e provided the DOMC is satisfied.

Sketch of proof Since, the first criterion of the DOMC is satisfied, the states $(x_i)_{1 \leq i \leq n}$ are nonlinearly observable (almost everywhere around x_e) and may be computed, by a unique way, in the $n - 1$ first rows of (15.5). So, in the last row of (15.5), the states vector x is determined, which allow to recover m almost everywhere around x_e if and only if $g_n \neq 0$.

We have proved how, under the DOMC, the inclusion of the message m in the transmitter's last dynamic will make it the last information recovered in the receiver. This, after the transmitter, states computation. Here after, we will describe the way to recover the message by a step-by-step delayed reconstructor.

15.3.4 The Deciphering Process

The synchronization between the transmitter and the receiver may be established, after the key exchange and thanks to the transfer of y . Now, the receiver is disposed to recover the message, under the DOMC, thanks to a step-by-step delayed reconstructor, which we describe in the following section.

15.3.4.1 Step-by-Step Delayed Reconstructor Design

The main idea of our technique is summed up in the following points:

1. The reconstructor dynamics $(\tilde{x}_i, \forall i \in \{2, \dots, n\})$ are computed step-by-step, one after the other.
2. In the reconstructor design, for a fixed iteration k :
 - The i th dynamic (i.e., $\tilde{x}_i$) is computed with $(k - i + 1)$ delays.
 - The $(i - 1)$ th dynamic is used to compute the $k - 1$ delayed value of the i th dynamic.
3. The n -time delayed value of u, u^{n-} is recovered in the last dynamic computed.

The step-by-step delayed reconstructor is given in different steps described by the algorithm hereafter. So, let us assume that system (15.5) at least, satisfied the DOMC almost everywhere around x_e .

15.3.4.2 The Deciphering Algorithm

15.3.4.2.1 Computation of $\tilde{x}_2^-$

In this first step, the only available information is y which verifies the equation: $y = f_1(y^-, \tilde{x}_2^-, p)$

So, according to the implicit function theorem, the dynamic $\tilde{x}_2^-$ is extracted from y if and only if $\partial f_1 / \partial x_2 \neq 0$ this condition holds in $\mathfrak{N}^n / \hat{S}_1$.

In this case, there exists a differentiable function F_1 such that $\tilde{x}_2^- = F_1(y^-, y, p)$.

When $x^- \in S_1$, $\tilde{x}_2^-$ take its last remembered value $\tilde{x}_2^- := (\tilde{x}_2)^{2-}$ in order to overcome this bifurcation.

This operation is performed for all the reconstructor dynamics:

15.3.4.2.2 Computation of $(\tilde{x}_{i+1})^{i-}$ for $2 \leq i \leq n - 1$

In the previous step (i.e., the i th step), the dynamic $(\tilde{x}_i)^{(i-1)-}$ was computed at the $(k - i + 1)$ th iteration, which allows to construct $(\tilde{x}_{i+1})^{i-}$ at the $(k - i)$ th iteration by using:

- The delayed dynamics $(\tilde{x}_2)^{i-}, \dots, (\tilde{x}_i)^{i-}$.
- The delayed output $(y)^{i-}$.
- The observed delayed dynamic:

$$(\tilde{x}_i)^{(i-1)-} = -f_i((y)^{i-}, \dots, (\tilde{x}_i)^{i-}, (\tilde{x}_{i+1})^{i-}, p)$$

So, when $x^{i-} \in \mathfrak{N}^n / S_i$ there exists a differentiable function F_i such that

$$(\tilde{x}_{i+1})^{i-} = F_i((y)^{i-}, \dots, (\tilde{x}_i)^{i-}, (\tilde{x}_i)^{(i-1)-}, p)$$

and for $x^{i-} \in S_i$ $(\tilde{x}_{i+1})^{i-} := (\tilde{x}_{i+1})^{(i+1)-}$. The isolation of such manifolds is the “key” of the reconstructor design.

15.3.4.2.3 Recovering of the information m^{n-}

After the computation of the last $(n - 1)$ -times delayed dynamic $(\tilde{x}_n)^{(n-1)-}$, in the last step, the information can be recovered with n delays:

$$m^{n-} = \frac{[(\tilde{x}_n)^{(n-1)-} - f_n((x_1)^{n-}, \dots, (\tilde{x}_{n-1})^{n-}, (\tilde{x}_n)^{n-}, p)]}{g_n}$$

All these steps are performed for each transmission iteration. Our delayed reconstructor is conceived such that, in each iteration, the information is extracted at the n th iteration upstream.

15.3.5 Observability Bifurcations

To tighten up the transmission security, we could either extend the bifurcations manifolds $(S_i)_{i \in I}$ or provoke them (see Remark 8).

To reach this goal, a so-called “extending functions” is implemented, for $i \in I$, in the transmitter corresponding dynamic, x_i :

$$\theta_i(x_1, \dots, x_{i+1}, p) = \begin{cases} 0 & \text{for } \left| \frac{\partial f_i}{\partial x_{i+1}} \right| \leq \alpha \\ x_i & \text{else} \end{cases}$$

These functions extend the singularities in a neighborhood of the bifurcations manifolds, delimited by a chosen threshold α , such a precision reduces or increases the new bifurcation region. Hence, during a possible decryption attempt, the interceptor ignores the existence of such bifurcation manifold (or at least the threshold value) and tries to observe a large unobservable manifold. By consequence, its decryption algorithm diverge quickly, loosing all possibility to extract the message.

REMARK 8

Unlike continuous-time systems, it is difficult to expect, for discrete-time systems if their trajectories jump on the bifurcations manifolds (if $\exists i$ such that, $\exists k \in \mathbb{N}$ for which $f_o^k(x, p) \in S_i$). Consequently, in the secured data transmission context, a function that provoke an observability bifurcation plays also the necessary role of function that extend it.

For example, the Rössler map presented in Chapter 9, which can be re-written under the following shape (with $f(x, p) := [f_1(x, p),$

$f_2(x, p), f_3(x, p)]^T$:

$$\begin{cases} x_1^+ = f_1(x, p) = a_1 x_1 (1 - x_1) + a_2 x_2 \\ x_2^+ = f_2(x, p) = b_1 \text{osg} \cdot (1 - b_4 x_3) \\ x_3^+ = f_3(x, p) = c_1 x_3 (1 - x_3) - c_2 (\text{osg} + 1) \end{cases} \quad (15.9)$$

possesses an observability bifurcation $\partial f_2 / \partial x_3 = \text{osg} = 0$. However, the simulation (performed for $N = 10^4$ iterations) results showed that $\nexists 0 \leq k \leq N$, for which $f_o^k(x, p) \in S_2$, in fact, $-1.4 < \text{osg} < -0.5$. So, in this case, the bifurcation may be provoked by injecting $\overline{\text{osg}}$ in x_2^+ and x_3^+ (we replace in 15.9 osg by $\overline{\text{osg}}$), in order to add to the security of the transmission.

$$\overline{\text{osg}} = \begin{cases} \text{osg} & \text{for } -0.7 < \text{osg} < -0.65 \\ 0 & \text{else} \end{cases}$$

The main idea of our approach: the design of a delayed observer (which we denote here as a reconstructor). This concept was presented in [5, 6, 10], where the authors used the $(i - N)$ -times ($N \in \mathbb{N}$) delayed values of the output y in order to observe the i th dynamic. Whereas the delayed reconstructor presented in this chapter take in account not only the delayed values of the output, but also the observed delayed dynamics. But the fundamental difference between both approaches is that the first one is based on a global observer, whereas the DCCIM has the worry to make a detailed local observation; by taking in account the exhibited observability bifurcations. Moreover, the main contribution of this technique is the opportunity to exploit these singularities, which represent a great asset for the cryptosystem.

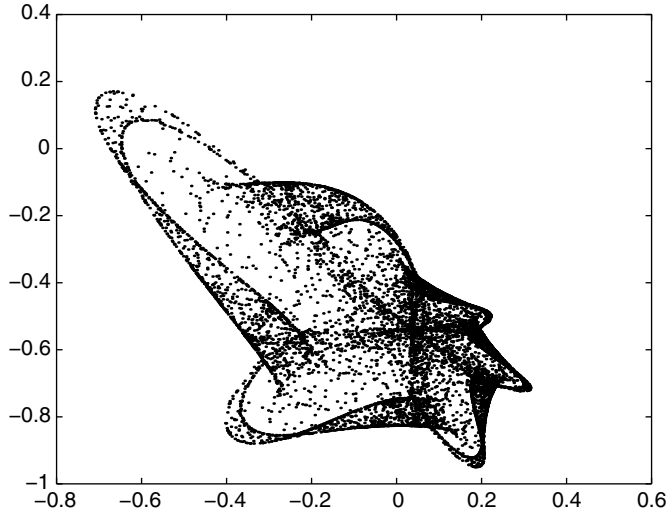
In the next section, we illustrate the DCCIM technique by an application example.

15.4 The Mandelbrot Map for the DCCIM

The ciphering process used in this transmission scheme is a chaotic generator widely studied in the literature: the Mandelbrot map (Figure 15.4). It is described in $\mathbb{R}^2$ by the following equations:

$$\begin{cases} x_1^+ = a + b + 2x_1 x_2 \\ x_2^+ = d + (a + c)x_2 + x_2^2 - x_1^2 + m \end{cases} \quad (15.10)$$

In this example, the private key is composed by the parameters: a and b . The constant d is fixed to 0.291.

**FIGURE 15.4**

Phase portrait of the Mandelbrot map for $a = 0.2$, $b = -0.7$, $c = 0.8$, and $d = 0.291$.

REMARK 9

System (15.10) is already under the discrete-time observability normal form given in Chapter 9.

REMARK 10

To preserve the chaotic behavior of the transmitter, the message amplitude should be adjusted with one of the Mandelbrot maps (i.e., 10^{-2}). To this end, a reducing parameter (depending on the message amplitude) may be introduced in the transmitter. This parameter should be manipulated carefully in the deciphering operation.

To exploit in the better sense the features of the Mandelbrot map for the ciphering process, a structural analysis of this one will be achieved.

15.4.1 Structural Analysis of the Mandelbrot Map

The bifurcations diagram of the map according to the parameter, c , of which does not depend the key, will be drawn first.

15.4.1.1 Bifurcations Diagram

A qualitative approach by the bifurcations diagram allows to determine the validity domain of the parameter c . Once the key is fixed, we draw this

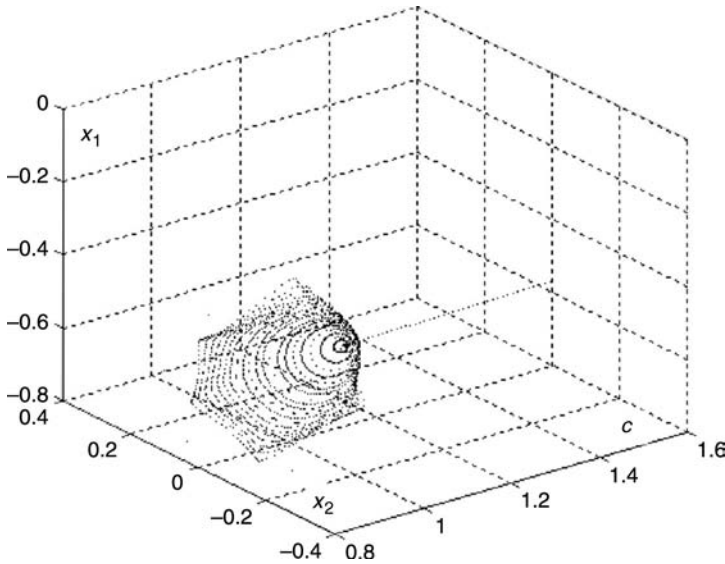


FIGURE 15.5
Three-dimensional bifurcations diagram.

diagram for different values of c (for well-chosen initials conditions). We summarize the main properties deduced as follows:

- The values of c for which the application trajectory is chaotic divide up, on intervals spread between 0.88 and 1.06 (see Figure 15.5).

REMARK 11

The bifurcations diagram is obtained as follows. For each value of c (with $0.8 \leq c \leq 1.6$), the Mandelbrot map is iterated 10 000 times without drawing (transient period) and we draw the 100 last points of the orbit.

- The application trajectory is periodic for some values of c such as 0.95 and 1 (they are pointed by an arrow in Figure 15.6). For $c = 0.95$, Figure 15.7, illustrates a periodic behavior of the Mandelbrot map, in some regions of the phase portrait (surrounded in this figure).
- For values of $c < 0.88$, the trajectory diverges quickly. However, for $c > 1.06$, the map possesses an attractor fixed point.

15.4.1.2 Lyapounov Exponents

When the parameter c is set up to 0.88, according to previous results, we should obtain at least one unstable Lyapounov exponent [computed from

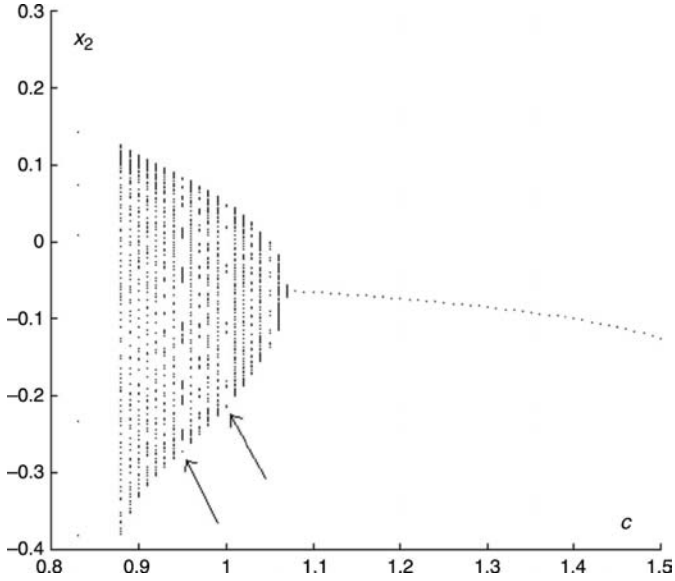


FIGURE 15.6
Bifurcations diagram.

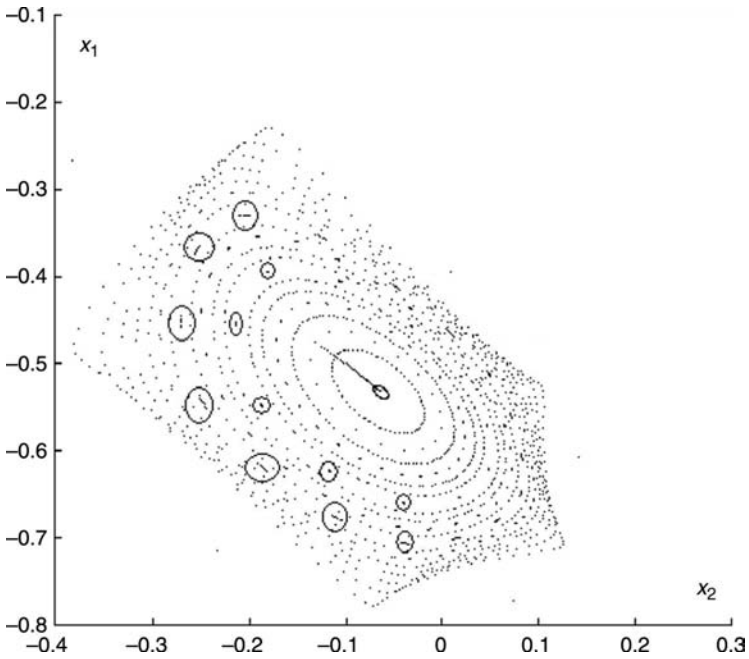


FIGURE 15.7
Mandelbrot map's phase portrait for $c = 0.95$.

the formula (15.2)]. Simulation results show that the Mandelbrot map has an hyperchaotic behavior for this value. In fact, both exponents are positive ($\lambda_1 = 0.1742$ and $\lambda_2 = 0.1721$). In contrast, both exponents are negative when $c = 1.5$. In this case, the phase portrait contracts and the trajectory converges to a fixed point.

15.4.1.3 Arnold's Tongue

The validity domain of c has been determined, it is now, necessary to establish accurately the validity domain of the key. This is justified on one hand by the worry of a correct ciphering. On the other hand, the aim of this evaluation is to estimate the difficulty of the exhaustive attack by the determination of the maximal region to cover. Provided the knowledge of the necessary time to cover a minimal region and the type of convergence of the algorithm used in this investigation [$O \ln(N)$, $O(N^2)$, ... with N the required number of iterations].

Matlab's simulations allow to estimate the Arnlod's tongue which corresponds to the couples (a, b) belonging to the domain:

$[-0.8, 0.25] \times [-1, 1]$ and given in Figure 15.8. These couples are those for which, at least one Lyapounov exponent is positive.

REMARK 12

A similar investigation has been carried out in order to estimate the validity domain of the initial conditions. So, for an optimal ciphering, we choose for $c = 0.8$,

$$x_1(0) = -0.2744, \quad x_2(0) = -0.452 \text{ and the key } p = (0.2, -0.7)^T.$$

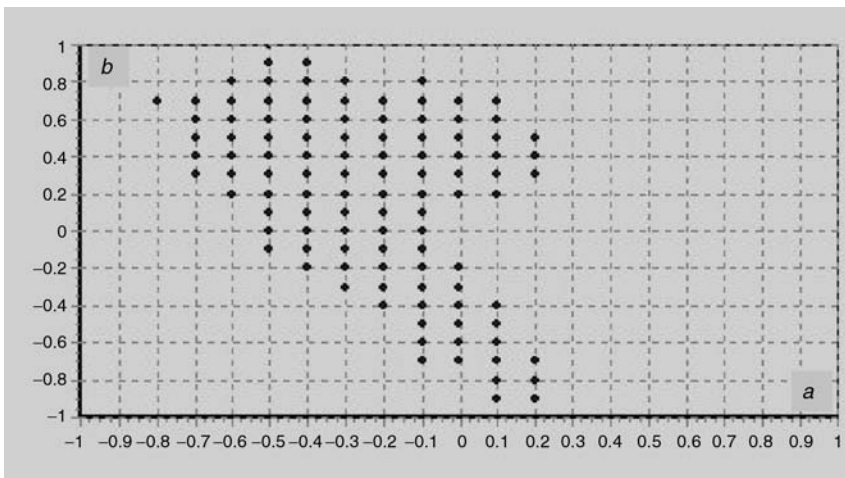


FIGURE 15.8

Arnold's tongue for the Mandelbrot map.

15.4.2 Message Deciphering

Before describing the deciphering process, we check whether system (15.10) satisfies the DOMC.

15.4.2.1 The DOMC for the Mandelbrot Application

15.4.2.1.1 DOMC First Criterion

The determinant of J is given by:

$$\det(J) = \begin{vmatrix} 1 & 0 \\ 2x_2 & 2y \end{vmatrix} = 2y$$

It ensues that $\det(J) \neq 0$ for $x \in \mathbb{R}^2/S_1$.

Hence, the matrix J is of rank 2 everywhere except in S_1 .

We deduce that system (15.10) is almost everywhere nonlinearly observable with one unobservable mode in the direction of x_2 . Thus, when $x \in S_1$, the observability bifurcation manifold:

$$S_1 = \{x \in \mathbb{R}^2 \mid y = 0\}$$

15.4.2.1.2 DOMC second criterion

In this example, $g_2 = 1 \neq 0$. So, the second criterion is verified⁵ and the message is recovered everywhere in $\mathbb{R}^2/S_1$.

15.4.2.2 The Step-by-Step Delayed Reconstructor

According to the algorithm performed in Section 15.3.4, the first step is to compute $\tilde{x}_2^-$ (the transmitter dynamic x_2^- reconstructed in the receiver). This dynamic may be observed thanks to the transmitter resonant terms, $x_2 y$. So,

$$\tilde{x}_2^- = \begin{cases} \frac{y - b - c}{2y^-} & \text{if } y^- \neq 0 \\ (\tilde{x}_2)^{2^-} & \text{else} \end{cases}$$

We recover the message with two delays, in the second step, as follows:

$$m^{2^-} = \tilde{x}_2^- - d - (a + c) (\tilde{x}_2)^{2^-} - \left((\tilde{x}_2)^{2^-} \right)^2 + (y^{2^-})^2$$

REMARK 13

In a ciphering–deciphering operation, in order to avoid the loss of information in the transient region, we add a temporal sample in the beginning of the message. In this example, two digital words are required.

⁵The product matrix vector $J^T \cdot g = (0, 0, 2y)^T$.

15.4.3 Bifurcation Extending Function

The efficiency of the observability bifurcations in a secured communication has been discussed in Section 15.3.5. In this section, we apply this concept for the Mandelbrot map.

Consequently, the observability manifold S_1 previously highlighted will be extended⁶ thanks to the function θ_1 . Here, $\partial f_1/\partial x_2 = 2x_1$. So,

$$\theta_1(x_1) = \begin{cases} 0 & \text{if } |x_1| \leq \alpha \\ x_1 & \text{else} \end{cases}$$

This function replaces the dynamic x_1 in the transmitter equations. Hence, when the singularity is detected, the value of x_1 will be replaced by zero. By consequence, the singularity domain will be extended to $[-\alpha, +\alpha] \times \Re$.

REMARK 14

The threshold is chosen with respect to the chaotic signal amplitude, for the Mandelbrot map, it should not exceed 10^{-2} .

However, this implementation may disturb the transmitter trajectory, which leads to an inaccurate deciphering in the receiver. The proposed solution is to interrupt the transmission when the bifurcation is detected. A fictitious message may be transmitted instead of the original one during the bifurcation phase.

15.5 A Ciphering–Deciphering Software

To experiment the CCMID technique in a secured transmission application, we have developed a software (in C language) for data ciphering and deciphering, such as pictures (Figure 15.9 and Figure 15.10), texts (Figure 15.11), as well as songs.

The difficulty that we met during the ciphering–deciphering operation is the considerable increase of the ciphered file size in comparison with the size of the original file (see Figure 15.12 and Figure 15.13).

This phenomenon is due to the dynamical feature of the cryptography by chaos. Let us take the example of a text file transmission. On the transmission line, the ciphered information is a real nonperiodic signal, it is represented by floating numbers sequences in double precision (8 bytes).

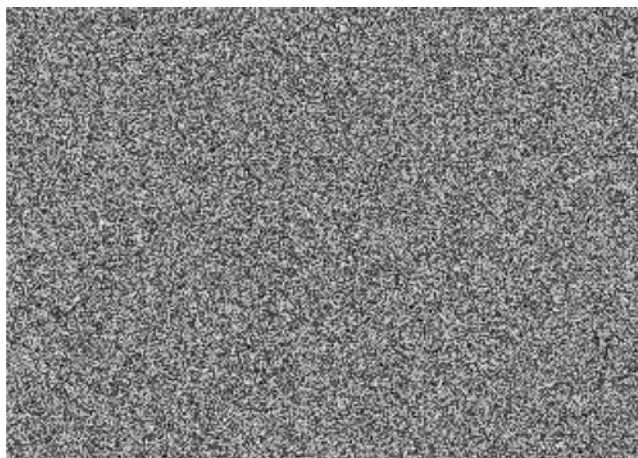
⁶Simulation results show that the Mandelbrot map trajectory reaches the bifurcation manifold S_1 . This allows to introduce an extending function. Otherwise, we should use a function to *provoke* the bifurcation.

**FIGURE 15.9**

The original and recovered picture.

Whereas the original message is constituted of characters sequences, encoded each one, on one byte.

Hence, each input character corresponds to eight output characters. This is why the ciphered text is eight times bigger than the original one.

**FIGURE 15.10**

The ciphered picture.

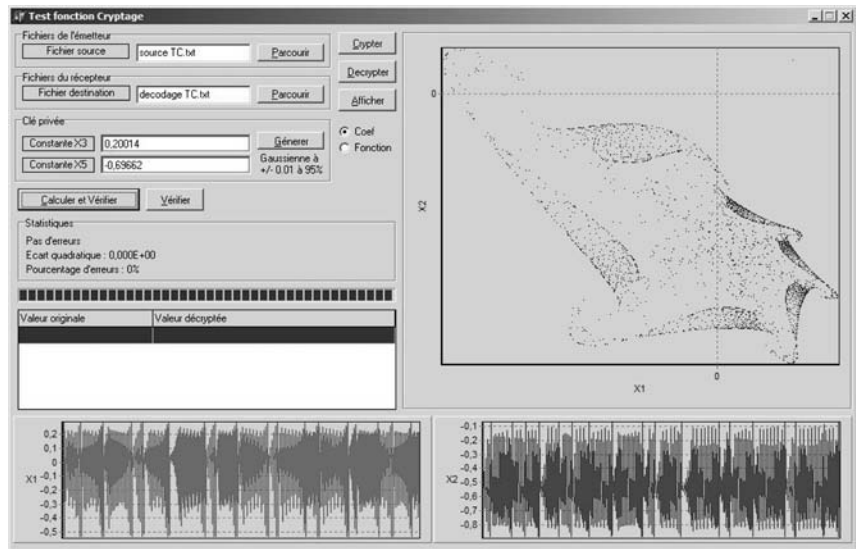


FIGURE 15.11
An example of cipheryng/decipheryng a text file by the CCMID.

Synchronisation de systèmes
chaotiques en temps discret :
Application à la cryptographie

FIGURE 15.12
Original text.

û2 ?i'î DÕ%½? ? 5¥Ð >zÍ]°@K3
†,àOhhr ½^6Æ}>, \$? %8 ”TMÄ-
Z6 Æ³ÜbP4ËÝ'Tr±YtU2ËªP. YÖ
d=½û612 öÏÖ4,ÉªÝ †Qâx ?švÝ
ÄCE¥ δYÄ01½\²Nbrö"Î ? }ÆjÎÉ
=?æÚ¹sJ-ën²i, çIâ!¼½úh Ac&³
Ä?ª8 ??-ûë ZCE{*>‡@sí/Šâp²ð
. ã TòU\$ < P}úÜ«â«Ä±LR ?\$úú
,º2%1ê><"ýü...«*E P GiWø?2,"Ì
ÎFαøÄ¥Ñ móy%ñè\i?÷ ?j æÄ|Û
z`?,eP” Óhh%@ Á? Ä\? =Üñ (‡
âF9`Gp¥ÖP ? 4OÑ"”ÍjCE”ÉSD²
ÛÁ†%V°U.α<'pÆWu2ð³jfi?.,? H

FIGURE 15.13
Cipheryng text (Figure 15.12) by the Mandelbrot map (in simple precision).

To avoid this problem, we have developed a specific application in order to reach the minimum ratio: (data in clear / ciphered data) allowing a correct ciphering and deciphering.

Currently, we are able to reach a minimum ratio equal to four. In other words, we cipher 2 bytes for 8 bytes read.

We can even reach a ratio equal to three which corresponds to two ciphered bytes for 6 bytes read. This operation is possible exclusively by using a standard floating type: *Real48* (in Delphi language).

15.6 Conclusions

The purpose of this chapter was to present a new algorithm of cryptography by chaos, in discrete time, using mathematical and automatic tools. This algorithm is in line with several experiences conducted through the world.

The reader may be interested to know that many works are being carried out actually, in relation with a great European project: the design of a ciphering broadband system exceeding gigabits per second. In particular using completely optic architectures [9].

The chaos made by complex dynamics would serve to generate random digital sequences much more preferable than the pseudo random computer sequences, used until now. One may enrich these ideas and imagine that the new notions of cryptography by chaos will serve to create private key for classical cryptographic algorithms. Effectively, algorithms founded on chaotic dynamics already exist in classical numerical cryptography and systems generating random numbers by chaotic dynamics were already proposed in quantique cryptography. For example, in March 2001, a ciphering software based on chaos theory appeared on the Internet. This system proposed by a Japanese society, announced a chaotic secrete key equivalent to a numerical key of length 1024 bits.

References

1. I. Belmouhoub, M. Djemaï, and J-P. Barbot, Cryptography by discrete-time hyperchaotic systems, in *Proceedings of IEEE CDC 2003*, Hawaï, USA, 2003.
2. J. Carrol and S. Martin, The automated cryptanalysis of substitution ciphers, *Cryptologia*, 10 (4), 193–209, 1986.
3. H. Dang-Vu and C. Delcarte, Bifurcations et chaos, une introduction à la dynamique contemporaine avec des programmes en Pascal, Fortran et Mathematica, *Ellipses*, 2000, (in french).

4. M. Henon, A two dimensional mapping with a strange attractor, *Commun. Math. Phys.*, 50, 66–77, 1976.
5. H.J.C. Huijberts, T. Lilge, and H. Nijmeijer, Control perspective on synchronization and the Takens–Aeyels–Sauer reconstruction theorem, *Phys. Rev. L.*, 59 (4), 1999.
6. H.J.C. Huijberts, T. Lilge, and H. Nijmeijer, Nonlinear discrete-time synchronization via extended observers, *Int. J. Bifurcation Chaos*, 11 (7), 1997–2006, 2001.
7. N. Koblitz, *A Course in Number Theory and Cryptography*, Springer Verlag, New York, USA, 1987.
8. H.J. Korsch and H.-J. Jodl, *Chaos. A Program Collection for the PC*, Springer-Verlag, Berlin, 2nd ed., 1998.
9. L. Larger and J.-P. Goedgebuer, Le chaos chiffant, *Pour Sci.*, 36, 2002, (in French).
10. T. Lilge, Nonlinear Discrete-Time observers for synchronization problems, *LNCIS 244, New Direction in nonlinear Observer Design*, 491–510, 1999.
11. D.D. Wheeler, Problem with chaotic cryptosystems, *Cryptologia*, 13, 243–250, 1989.
12. S. Wiggins, *An introduction to applied nonlinear dynamical systems and chaos*, Springer, 1990.

Appendix A

On Ergodic Theory of Chaos

CONTENTS

A.1 Introduction	553
A.2 Theoretical Background	553
A.3 Conclusions	558
References	558

A.1 Introduction

With regard to both the mathematical description (as sets of differential equations, discrete maps, etc.) and their (disordered-like) aperiodic behavior, chaotic systems can be viewed as belonging to a particular class located at a virtual crossroad between determinism and randomness. As previously pointed out, numerous chaos control methodologies (such as the OGY method, the Pyragas method, the local H_∞ control method, etc.) rely directly or implicitly on ergodic properties of chaotic systems. In addition to the deterministic viewpoint considered in this book, this short appendix aims at briefly introducing some probabilistic considerations of chaotic motions, through the presentation of some important notions and results intrinsic to the ergodic theory.

A.2 Theoretical Background

According to the purpose of dealing with qualitative and quantitative properties of dynamical systems behaviors, the ergodic theory makes

extensive use of the concepts and formalism of the measure theory [7]. Let us introduce the formal notion of *measure*.¹

For this purpose, consider a space X of all possible states² $x \in X \subseteq \mathbb{R}^n$ related to the iteration k of a (possibly nonlinear) transformation T_k intrinsic to a given discrete time dynamical (possibly chaotic) system. Also consider, over the domain X , a σ -algebra (or σ -field) $\mathcal{S}$, that is, a nonempty collection of subsets of X such that the following hold:

1. The empty set $\{\emptyset\}$ is in $\mathcal{S}$
2. $A \in \mathcal{S}$ implies that its complement $A^c \in \mathcal{S}$
3. If $\{A_i\}_{i=1}^\infty$ is a sequence of elements in $\mathcal{S}$ then $\bigcup_{i=1}^\infty A_i \in \mathcal{S}$

DEFINITION 1 A measure is a (nonnegative) function $\mu : \mathcal{S} \rightarrow \mathbb{R}$ satisfying the following properties:

- $\mu(\{\emptyset\}) = 0$ (where $\{\emptyset\}$ denotes the empty set)
- $\mu(A) \geq 0$ for all $A \in \mathcal{S}$
- For a countable sequence of disjoint sets $\{A_i\}_{i=1}^k$ with $A_i \in \mathcal{S}$ and $k \in \mathbb{N} \cup \infty$, we have the following *countable additive condition*,³

$$\mu \left(\bigcup_{i=1}^k A_i \right) = \sum_{i=1}^k \mu(A_i) \quad (\text{A.1})$$

If, in addition, $\mu(X) = 1$, then μ is a *probability measure*.

According to this general definition, there exists an infinite number of possible measures (e.g., the Dirac measure, the Borel measure, the counting measure, the Lebesgue measure, etc.) for a given dynamical system. However, within the context of chaotic systems, some measures such as invariant and ergodic ones are of particular interest. Thus, in what follows, we define such measures via the consideration of transformations [2].

For this purpose, consider two *probability spaces* $(X_{i=1,2}, \mathcal{S}_{i=1,2}, \mu_{i=1,2})$ where $(X_{i=1,2}, \mathcal{S}_{i=1,2})$ are *measurable spaces*, $\mu_{i=1,2}$ are *probability measures*, and σ -algebras $\mathcal{S}_{i=1,2}$ are families of measurable subsets (*events*) of $X_{i=1,2}$.

¹A measure is a function which assigns lengths, volumes, or probabilities to subsets of a given set.

²From a probabilistic or statistical viewpoint, X can be viewed as a *sample space*, that is, a set of *outcomes*, possibly infinite in number.

³A finitely additive measure has the same definition except that $\mathcal{S}$ is only required to be an algebra and the additive property is only required to hold for finite unions.

DEFINITION 2 Let $(X_1, \mathcal{S}_1, \mu_1)$ and $(X_2, \mathcal{S}_2, \mu_2)$ be two probability spaces. A transformation $T : X_1 \rightarrow X_2$ is said to be measurable if $T^{-1}(A) \in \mathcal{S}_1$ for every $A \in \mathcal{S}_2$. If, in addition, $\mu_1(T^{-1}(A)) = \mu_2(A)$ for every measurable set $A \in \mathcal{S}_2$, then T is said to be a measure-preserving transformation. In addition, if $(X_1, \mathcal{S}_1, \mu_1) = (X_2, \mathcal{S}_2, \mu_2)$ then T is a measure-preserving endomorphism.

Remark: With respect to this definition, if $\mu(T^{-1}(A)) = \mu(A)$ then μ is said to be T -invariant (i.e., invariant under T).

Now, through the consideration of ergodic measures, Definition 2 can be related to ergodic transformations as follows.

DEFINITION 3 A measurable-preserving transformation $T : X \rightarrow X$ on a probability space $(X, \mathcal{S}, \mu)$ is said to be ergodic if for every measurable set $A \in \mathcal{S}$ with $T^{-1}(A) = A$, we have $\mu(A) = 0$ or $\mu(A) = 1$.

In such a case, μ is said to be T -ergodic.

This last definition states that T acts on almost all sets all over the space except sets of measure zero and the entire space. In other words, a transformation is ergodic if (and only if) the orbit of almost every point visit each set of positive measure.

Chaotic systems are often referred to as processes having a *mixing property*⁴ [4]. We introduce here some definitions related to this property and some links with ergodic transformations (see Lemma 1 and Lemma 2).

DEFINITION 4 (weak-mixing) Let $T : X \rightarrow X$ be a measure preserving transformation on a probability space $(X, \mathcal{S}, \mu)$. Then, T is weak-mixing if for any subsets $A, B \in \mathcal{S}$ we have,

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{j=1}^N \left| \mu(T^{-j}A \cap B) - \mu(A)\mu(B) \right| = 0 \quad (\text{A.2})$$

LEMMA 1

If a transformation $T : X \rightarrow X$ on a probability space $(X, \mathcal{S}, \mu)$ is weak-mixing then it is necessarily ergodic.⁵

DEFINITION 5 (strong-mixing) Let $T : X \rightarrow X$ be a measure preserving transformation on a probability space $(X, \mathcal{S}, \mu)$. Then, T is strong-mixing if for any subsets $A, B \in \mathcal{S}$, we have,

$$\lim_{j \rightarrow +\infty} \mu(T^{-j}A \cap B) = \mu(A)\mu(B) \quad (\text{A.3})$$

⁴More precisely, an exponentially fast mixing property.

⁵But the converse is not true.

LEMMA 2

If a transformation $T : X \rightarrow X$ on a probability space $(X, \mathcal{S}, \mu)$ is strong-mixing then it is necessarily weak-mixing (and therefore ergodic).

Now, with such definitions in mind, we can introduce a fundamental result, known as Poincaré recurrence theorem and express it as follows.

THEOREM 1 (Poincaré recurrence theorem)

Let $T : X \rightarrow X$ be a measurable transformation on a probability space $(X, \mathcal{S}, \mu)$ and μ a T -invariant finite measure on X . If $A \subset X$ is a measurable subset (i.e., $A \in \mathcal{S}$) with positive measure (i.e., $\mu(A) > 0$) then

$$\text{card} \{n \in \mathbb{N} : T^n x \in A\} = \infty \quad (\text{A.4})$$

for μ -almost every point $x \in A$.

This theorem states that any dynamical system preserving a finite measure, exhibits a (nontrivial) recurrence in any set $A \in \mathcal{S}$ with positive measure, so that the orbit of almost every point in A returns infinitely often to A under the iteration of T .

With regard to some (local) chaos control methods (such as the OGY method and the Pyragas technique), Poincaré's recurrence theorem appears to be of crucial importance, as it guarantees that trajectories of a chaotic systems (with measure-preserving transformations) remain in a bounded region of the state space (without entering within regions with nonpositive measure that are related to the attractor) [3]. However, note that Poincaré's recurrence theorem is essentially qualitative as it gives no information about the frequency with which each trajectory visits a given set. This lack of information was overcome by Birkhoff [1] and von Neumann [9] who established independently the first version of the ergodic theorem. Before introducing such a fundamental result, the first step in quantifying some aspects of the recurrence theorem is presented.

Let $n_A : A \rightarrow \mathbb{Z}^+ \cup \{+\infty\}$ be the first return time⁶ (i.e., $n_A(x) > 0$ is the smallest value for which $T^{n_A(x)} x \in A$). Then, if the measure μ is T -ergodic then the *average* return time to A can be defined according to the following theorem [5].

THEOREM 2 (Kac's return time theorem)

Let $T : X \rightarrow X$ be an ergodic transformation on a probability space $(X, \mathcal{S}, \mu)$. Let $A \in \mathcal{S}$ have $\mu(A) > 0$ then we define the return time function $n_A : A \rightarrow \mathbb{Z}^+ \cup \{+\infty\}$ (which is finite, almost everywhere). The average return time (with respect to the induced probability measure μ_A) is

$$\int_A n_A(x) d\mu_A(x) = \frac{1}{\mu(A)} \quad (\text{A.5})$$

⁶With respect to Theorem 1, n_A is finite almost everywhere.

REMARK 1

For a given (discrete time) ergodic dynamical system, Kac's theorem gives some quantitative informations about the (statistical) duration to wait to return within any small neighborhood of an original position in the phase space. However, it is worth mentioning that, due to the sensitivity to initial conditions and exponential instability, this return time property cannot be related to "periodical motion" *stricto sensu* [8].

Finally, to complete this quantitative insight of ergodic properties, express the seminal theorem of Birkhoff.⁷

THEOREM 3 (Birkhoff's ergodic theorem)

Let $(X, \mathcal{S}, \mu)$ be a probability space with μ , an ergodic measure and $\mathcal{S}$, a σ -algebra. Moreover, let $f : X \rightarrow \mathbb{R}$ be a real-valued measurable function. Then for almost all $x \in X$ we have,

$$\frac{1}{N} \sum_{j=1}^N f \circ T^j(x) \rightarrow \int f \, d\mu \quad (\text{A.6})$$

as $N \rightarrow \infty$.

The left-hand side of (A.6) just says how often the orbit of x (i.e., the points x, Tx, T^2x , etc.) lies in A and the right-hand side is just the measure of A . Thus, for an ergodic endomorphism, we have "space-averages = time-averages almost everywhere."⁸ Nevertheless, this theorem considers only one aspect of the quantitative behavior of recurrence. In particular, it gives no information about the rate at which a given trajectory returns arbitrarily close to itself [6]. The following corollary shows that the Birkhoff theorem can also give a quantitative version of the Poincaré recurrence theorem.

COROLLARY 1

Let $(X, \mathcal{S}, \mu)$ be a probability space, and assume that the transformation $T : X \rightarrow X$ preserves μ and is ergodic. The proportion of time spent by almost all points in a subset $A \in \mathcal{S}$ is given by its measure $\mu(A)$, that is,

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \text{card}\{0 \leq n \leq N-1 : T^n x \in A\} = \mu(A) \quad (\text{A.7})$$

for almost all points $x \in X$.

⁷In its version involving ergodic measure.

⁸Note that this notion of time-averages is important as it implies that transients become irrelevant.

REMARK 2

When dealing with (local) control methodologies for chaotic systems, Poincaré's recurrence theorem and Birkhoff's ergodic theorem are of crucial interest. Indeed, these results provide some theoretical tools to focus on some chaotic motion properties such as the filling of some bounded regions of the state space (with respect to the attractor characteristics). Then, such results can serve to ensure that a (chaotic) system trajectory will always visit an expected (small) neighborhood of a state where the control law has to start acting, and that the trajectory will also return, in finite time, to this neighborhood in case of lost of effectiveness of the control law (leading the system evolution to become once more wandering).

A.3 Conclusions

As briefly presented in this appendix, the ergodic theory aims at studying the dynamics of general measurable maps (i.e., transformations) on general measure spaces, by means of concepts and formalism intrinsic to the measure theory [10–12]. In particular, this theoretical framework provides numerous results to understand and describe some probabilistic and statistical properties of some classes of dynamical systems (including chaotic ones), within the context of measure-preserving and ergodic transformations (or invariant and ergodic measures). Such results appear to be of crucial importance for dealing with global behavior of chaotic systems. Indeed, despite the unpredictability of long-term behaviors of such systems, these results provide some probabilities (or even certitudes) of occurrence of some behaviors or events, which may or may not enable the development of some chaos control methodologies within a deterministic context.

References

1. G.D. Birkhoff, Proof of the ergodic theorem, *Proc. Natl. Acad. Sci. USA*, 17, 656–660, 1931.
2. G.D. Birkhoff, *Collected Mathematical Papers*, vol. 3, Dover, N.Y., 1960.
3. J.P. Eckmann and D. Ruelle, Ergodic Theory of Chaos and Strange Attractors, *Rev. Mod. Phys.*, 57 (3), 617–656, 1985.
4. J.L. Jensen, Chaotic dynamical systems with a view towards statistics: a review, in *Networks and Chaos, Statistical and Probabilistic Aspects*, O.E. Barndorff-Nielsen, J.L. Jensen, and W.S. Kendall, Eds., Chapman and Hall, London, 1993.

5. M. Kac, On the notion of recurrence in discrete stochastic processes, *Bull. Am. Math. Soc.*, 53, 1002–1010, 1947.
6. A.G. Kachurovskii, The rate of convergence in ergodic theorem, *Russ. Math. Survey*, 51, 653–705, 1996.
7. A. Katok and B. Hasselblatt, *Introduction to Modern Theory of Dynamical Systems*, Cambridge University Press, 1995.
8. R. Mañé, *Ergodic Theory and Differentiable Dynamics*, Springer-Verlag, 1987.
9. J. von Neumann, Proof of the quasi-ergodic hypothesis, *Proc. Natl. Acad. Sci. USA*, 18, 70–82, 1932.
10. Ya.G. Sinai, *Introduction to Ergodic Theory*, Princeton University Press, 1977.
11. Ya.G. Sinai, *Topics in Ergodic Theory*, Princeton University Press, 1994.
12. P. Walters, *An Introduction to Ergodic Theory*, Graduate Text in Mathematics, vol. 79, Springer, New York, 2000.

Index

- σ -algebra, 554
- absolutely continuous, 52
- Adaptive Adjustment Mechanism (AAM), 315
- addition method, 393
- approximates
 - Euler, 56
 - Picard-Lindelöf, 57
- attractivity, 70
- attractor, 266
 - strange, 10, 60, 96, 265
- autonomous, 58

- Bernoulli relation, 47
- bifurcation, 4, 88, 266
 - codimension, 92
 - diagram, 90
 - fork, 94
 - Hopf, 91, 95, 118, 122, 123, 483, 489
 - observability, 377
 - subcritical, 92, 123, 129
 - supercritical, 123
 - transcritical, 93
 - value, 90
- Brunovsky form, 349

- capacity, 97
- Carathéodory, 55
- Cauchy
 - problem, 52
- Center manifold, 102
- chaos, 60, 96
 - H_∞ -control, 302
 - adaptive control, 304
 - control, 294
 - control: OGY method, 294
 - control: Pyragas method, 299
 - hyperchaos control, 312
 - sliding mode control, 306
- chemistry, 47
- Chua's circuit, 505

- Code Division Multiple Access (CDMA), 454
- codimension, 129
- commutation of two vector flows, 62
- complete
 - vector field, 61
- conjugacy, 79
- controllable
 - linearly, 348
- convex hull, 326

- Delay coordinate vector, 298
- Dependence of the initial conditions for an ODE, 58
- dimension
 - Embedding, 298
 - fractal, 97
 - Hausdorff, 97
 - Liapunov, 97
 - Rényi, 97
- Discrete-time Cryptography by Chaos based on the Inclusion Method (DCCIM), 528
- Discrete-Time Hyperchaotic-Cryptography by Inclusion Method, 399
- distance
 - Whitney - (or $C1$), 79
- domain
 - stability, 70
- Dunford splitting, 76
- dynamics
 - fast, 268
 - slow, 268

- eigenvalues, 119, 130
- eigenvectors, 115, 116
- equation
 - Hindmarsh and Rose, 280
 - ball and beam, 366
 - Chua's circuit, 48, 276, 393
 - Euler-Lagrange, 49
 - Liénard, 48
 - logistic, 47, 53

- equation (*Contd.*)
 - Lorenz, 117, 373
 - Moore-Greitzer model of an engine
 - compressor, 370
 - nonlinear delay differential - (NLDDE), 458
 - ordinary differential, 5, 46
 - pendulum, 50
 - Rössler, 96, 129
 - Van Der Pol, 48, 66, 114
 - Volterra-Lotka, 63
- equilibrium
 - point, 62
 - branch, 89
 - degenerated, 65
 - point, 4, 102
 - saddle, 92
 - sink, 135
 - source, 135
- ergodic measure, 555
- ergodic transformation, 555
- Euler approximates, 56
- exponential stability, 74
- Extended Time-Delayed
 - AutoSynchronization method (ETDAS), 301
- field-oriented control (FOC), 481
- fixed point, 4, 102
- Floquet multipliers, 301
- Floquet theory, 60
- flow, semi-flow, 61
- Gronwall lemma, 58
- hyperbolic point, 65
- hyperchaotic system, 312, 399
- Ikeda Ring Cavity, 461
- inclusion method, 393
- initial condition, 52
- integral manifold approach, 263, 272
- invariant measure, 555
- jacobian, 57
 - matrix, 65
- Jordan
 - curve, 66
- Kazantzis-Kravaris Partial Differential Equations, (KK PDE), 411, 412
- Kotelianskii criterion, 78
- lagrangian, 49
- Liapunov
 - dimension, 97
 - exponent, 96, 98
 - first Liapunov method, 81
 - number, 123, 129
 - Parameter Dependent Liapunov Functions, 325
 - Parameter Dependent Liapunov Functions (PDLF), 329
- Lie bracket, 62
- limit cycle, 60, 121
- limit layer, 270
- Linear Matrix Inequalities (LMI), 329, 333
- linear parameter varying systems (LPV), 327
- linearity, 59
- lipschitzian function, 57
- Lyapunov
 - function, 378
- manifold
 - differentiable, 51
 - fast, 274
 - integral, 267
 - local unobservability, 391
 - local, stable, unstable, 82
 - slow, 274
- map
 - Burgers, 397, 529
 - Controlled Poincaré, 296
 - Mandelbrot, 541
 - Myrberg's, 9
 - Rössler, 399, 540
- matrix exponential, 75
- maximum solution, 57
- measurable, 52
- measurable space, 554
- measure-preserving
 - transformation, 555
- mixing property, 555
- motion, 53
- Neimark curve, 18
- node, 92
- normal form, 130
 - ball and beam, 367
 - Bogdanov, 107
 - controllable, 352
 - discrete-time observability, 390
 - Guckenheimer-Holmes, 132

- Lorenz, 372
- Moore-Greitzer model of an engine
 - compressor, 369
- observability, 387
- Poincaré-Dulac, 104
- Poincaré, 349
- Takens, 107
- with controllable and uncontrollable parts, 356
- observability
 - bifurcation, 378
 - Discrete-time Observability
 - Matching Condition (DOMC), 393
 - matching condition (OMC), 392
 - normal form, 387
- observable
 - weakly locally, 382
- observer, 411
 - Kreisselmeier and Engel, 414
 - unknown input, 335
- ODE (ordinary differential equations), 5, 46
 - Autonomous linear, 59
 - explicit, 51
 - implicit, 50
- OGY region, 294
- orbit, 53
 - closed, 66
 - heteroclinic, 66, 266
 - homoclinic, 66, 266
 - periodic, 65
- order, 50
- ordinary differential equation
 - autonomous- (or stationary), 59
 - linear non autonomous, non stationary, 60
- Parameter dependence (for an ODE), 88
- Park model, 49
- Peano, 55
- periodic
 - function, 60
 - orbit, 65
 - vector field, 59
- phase
 - portrait, 53
 - space, 51
- Picard-Lindelöf approximates, 57
- Poincaré
 - section, 86
 - section method, 295
- polytope, 326
- polytopic, 327
 - decomposition, 327
- observer, 332
- probability
 - measure, 554
 - space, 554
- problem
 - boundary, 52
 - Cauchy, 52
- QEMOI
 - see Quadratically Equivalent Modulo an Output Injection
- QEMOI for a discrete time system, 388
- Quadratically Equivalent Modulo an Output Injection (QEMOI), 384
- qualitative methods, 4
- rank-one image, 11
- repeller
 - strange, 10
- resonance, 115
- robustness, 79
- Rössler attractor, 96
- set
 - attractive, 71
 - invariant, 68
- singular perturbation, 263
- slow model, 270
- solution, 52
 - existence, 53
 - maximum, 57
- space
 - extended state, 53
 - of motion, 53
 - state, 51
 - tangent, 51
- stability
 - asymptotic, 72
 - domain, 69
 - global, local, 70
 - Liapunov, 69
 - poly-quadratic, 329
 - uniform, 69
- standard form, 268
- state vector, 51
- step by step sliding mode observer, 395

- step-by-step delayed reconstructor, 401
- stepper motor, 49
- strong-mixing property, 555
- structural stability, 5, 266
- structurally stable, 80
- synchronization, 332, 378
 - Input Independent Global, 336
- syntonization, 378
- theorem
 - Birkhoff's ergodic, 557
 - Carathéodory, 55
 - center manifold, 83
 - Dependence of the initial conditions, 58
 - Gröbner, 269
 - Hartman-Grobman, 82
 - Henry and Carr, 84
 - Kac's return time, 556
 - LaSalle, 379
 - Peano, 55
 - Poincaré recurrence, 556
 - Poincaré-Dulac, 104
 - stable manifold, 82
 - Tikhonov, 272
- time variable, 51
- Time-Delayed AutoSynchronization
 - method (TDAS), 300
- topological equivalence, 80
- trajectory, 53
- universal unfolding, 130, 132
- Unstable Periodic Orbit (UPO), 291
- vector field, 51
- weak-mixing property, 555
- Whitney distance, 79

Chaotic behavior arises in a variety of control settings. In some cases, it is beneficial to remove this behavior; in others, introducing or taking advantage of the existing chaotic components can be useful for example in cryptography. **Chaos in Automatic Control** surveys the latest methods for inserting, taking advantage of, or removing chaos in a variety of applications. This book supplies the theoretical and pedagogical basis of chaos in control systems along with new concepts and recent developments in the field.

Presented in three parts, the book examines open-loop analysis, closed-loop control, and applications of chaos in control systems. The first section builds a background in the mathematics of ordinary differential and difference equations on which the remainder of the book is based. It includes an introductory chapter by Christian Mira, a pioneer in chaos research. The next section explores solutions to problems arising in observation and control of closed-loop chaotic control systems. These include model-independent control methods, strategies such as H-infinity and sliding modes, polytopic observers, normal forms using homogeneous transformations, and observability normal forms. The final section explores applications in wireless transmission, optics, power electronics, and cryptography.

Features

- Provides an integrated reference to the analysis, design, and applications of chaos in control systems
- Takes a deterministic approach based on ordinary differential and difference equations
- Introduces chaos theory progressively from bifurcation theory and singular perturbation to normal forms
- Examines synchronization, observability bifurcation, chaotic and hyper-chaotic control, and control bifurcation
- Includes more than 130 figures and 1140 equations to illustrate the concepts

Chaos in Automatic Control distills the latest thinking in chaos while relating it to the most recent developments and applications in control. It serves as a platform for developing more robust, autonomous, intelligent, and adaptive systems.



Taylor & Francis

Taylor & Francis Group

A CRC PRESS BOOK

www.taylorandfrancisgroup.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487

270 Madison Avenue
New York, NY 10016

2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

DK3143

ISBN 0-8247-2653-7

90000



9 780824 726539